Time Dependent Stochastic mRNA and Protein Synthesis in Piecewise-deterministic Models of Gene Networks.

Guilherme C.P. Innocentini[1], Arran Hodgkinson[2] and Ovidiu Radulescu[2]

[1]Instituto de Física, Universidade Federal de Uberlândia, Brazil

[2]DIMNP, UMR CNRS 5235 - University of Montpellier, Montpellier, France

Abstract We discuss piecewise-deterministic approximations of gene networks dynamics. These approximations capture in a simple way the stochasticity of gene expression and the propagation of expression noise in networks and circuits. By using partial omega expansions, piecewise deterministic approximations can be formally derived from the more commonly used Markov pure jump processes (chemical master equation). We are interested in time dependent multivariate distributions that describe the stochastic dynamics of the gene networks. This problem is difficult even in the simplified framework of piecewise-determinisitic processes. We consider three methods to compute these distributions: the direct Monte-Carlo, the numerical integration of the Liouville-master equation and the push-forward method. This approach is applied to multivariate fluctuations of gene expression, generated by gene circuits. We find that stochastic fluctuations of the proteome and much less those of the transcriptome can discriminate between various circuit topologies.

1 Introduction

One of the greatest problems of molecular biology is how single undifferentiated cells give rise to many different cells types, all being genetically identical yet performing different functions. Since the pioneering work of Jacob and Monod [24] it is known that this multiplicity of behaviours is possible because the protein production depends not only on the existence of a gene but also on the quantities of regulatory molecules that can change with the cell type and environmental conditions. The protein production takes place in two steps [2]. First, during a process called transcription, the genetic information from DNA is copied on the messenger RNA (mRNA). Then, during a process called translation, the mRNA is used as template for protein production by the ribosomes. The amount of transcribed mRNA and translated protein, namely the gene expression, can vary from one cell to another for various reasons. One of these reasons is that the gene expression is not a property of a single gene but is a property of a set of interacting genes; a gene network [26]. As per usual in complex systems, the whole is different from the sum of its parts. In this case, a gene network can have many different stable expression levels that correspond to different network attractors [50]. Another reason for expression variability, is the fact that transcription and translation are stochastic [32, 11, 49]. A single cell transcription is often intermittent, periods of strong mRNA production being followed by periods of silence when transcription is stopped [49, 40, 6, 48]. The periods of inactivity may correspond to paused RNA polymerase, incompletely formed activation complexes, transient presence of regulatory proteins and complexes on the DNA, or to transitions between open and condensed chromatin [36]. The durations of the periods of activity and inactivity are random. Furthermore, snapshots of the cell population at various times show a heterogeneous picture in which a significant proportion of the cell

population express much lower or much higher than the average protein levels [11, 12]. The observed distributions of the gene expression can deviate significantly from a Gaussian by having skewness [12], heavy tails or multimodality [36].

The gene expression stochasticity can have important biological consequences. In infections and tumours, part of pathogen populations (bacteria, viruses, cancer) can escape drug treatment when they adopt behaviours very different from the average; for instance they may stop growth [10, 3], remodel their metabolism [19], or stop absorbing, or expel drugs [18]. Latency phenomena [43] can be responsible for resistance to treatment and relapse of the disease once the treatment is stopped.

Gene expression stochasticity is also important in synthetic biology. Synthetic biology devices are supposed to have a well defined biological function for given combinations of the input conditions. Therefore, in the absence of error-correction, the precision of biological devices depends critically on the amount of stochasticity [41, 35].

For all these reasons, one needs mathematical tools for predicting quantitatively the amplitude and distribution of gene expression fluctuations in gene networks.

Models of stochastic gene networks represent molecular interactions as discrete events (biochemical reactions) separated by random waiting times. This modeling approach, introduced by Delbruck [1] and further developed by Renyi [44] and Bartholomay [4] covers practically all the aspects of gene/gene interactions but is mathematically and computationally challenging. Indeed, the underlying models are continuous time Markov processes with an infinite or extremely large number of states. The corresponding master equation can be solved exactly only in a limited number of cases, corresponding to single genes

without or with much simplified feed-back control [37, 20, 21, 42]. The Monte-Carlo simulation algorithms (Gillespie algorithm [15]) can be very inefficient for simulating the full process, as the number of individual reactions that have to be simulated for a significant change of the system's state can be tremendously large.

Several types of approximation were used in order to simplify stochastic biochemical networks in order to reduce their simulation time and to facilitate their analysis. These simplifications were possible because the stochastic gene networks have heterogeneous variables and multiple time scales [39]. This heterogeneity comes from the fact that some variables X_D such as DNA/regulatory proteins and complexes/polymerase states are discrete and other variables X_C such as protein and mRNA copy numbers are continuous. The species dichotomy leads to a partition of the biochemical reactions. Although literature provides a number of different ways for reaction partitioning, the four set partition R_D, R_C, R_{DC}, R_{CD} (Figure 1) seems to us quite natural. A very similar partitioning was used for rigourously justified approximations in [8, 7].

Two main timescales have to be taken into account in order to find the appropriate approximation. As shown in the Figure 1, the discrete variables switch between a number of discrete states. The characteristic time of this process was called switching time, τ_S [39]. The trajectories of continuous variables are smooth only in the average, but for these variables, the average is a good approximation. The characteristic time of fluctuations of continuous species around their average was named discreteness time, τ_D [39]. The discreteness time scales like $1/N$ where N is the copy number of the continuous species.

There are two main classes of approximated models (Figure 2).

- Deterministic (ordinary differential equations, ODEs) or diffusion (Fokker-Planck) approximations

can be applied when the transitions between discrete states are fast and the continuous species have large copy numbers, thus when τ_S, τ_D are both small. Then one obtains the deterministic or diffusion limit by applying the law of large numbers or the central limit theorem to Markov processes [28, 29]. This result can be obtained heuristically by the Ω expansion, well known in physics [51]. If the discreteness time is larger than the switching time, similar approximated models can be obtained by using averaging [8, 7]. In both cases the discrete states are homogenized and the approximated model has only continuous variables.

- Piecewise-determinisitic processes (PDP) or hybrid diffusion approximations can be obtained when the discreteness time is small, but $\tau_S >> \tau_D$. In these approximations there are two types of variables. The dynamics of the continuous variables is described by ODE or stochastic differential equations (SDE), whereas the discrete variables dynamics is described by Markov chains. These approximations were first obtained heuristically by using partial Ω expansions of the master equation [17, 38, 8] and then justified rigorously by using generators and measure theory in [7]. Finally, diffusion approximation was applied to PDP in the limit when the switching time is small, to obtain again deterministic and Fokker-Planck approximations [39].

In this paper we consider the situation when the swiching time is relatively large. In biology, this corresponds to the so-called "random telegraph"[30], "bursting" [40, 12], or "multi-scale bursting" [48] fluctuations. Simply speaking there are some variables (DNA, regulatory complexes and/or RNA polymerase states) that have ON-OFF or multiple states Markov chain dynamics. The discrete variables control the ODE dynamics of the continuous variables (mRNA, proteins). The underlying approximation is piecewise-

deterministic, because the ODEs change each time when the discrete variables perform a transition (see Figure 1).

In our PDP models of gene networks each gene promoter is described as a finite state Markov chain. The transition rates between states of the promoter depend on the expression levels of the regulating genes. The promoter triggers synthesis of gene products with intensities depending on its state. We represent the state of a gene network as a random vector whose components indicate the promoter state and the copy numbers of all the mRNAs and proteins in the network. The dynamics of the gene network will be represented as the time-dependent multivariate distribution of this random vector. These distributions contain information about the randomness of each gene but also about correlation among genes and promoter states. We present three methods to compute time dependent multivariate distributions for such models: the direct simulation of the PDP process (the Monte-Carlo method), the numerical solutions of the Liouville-master equation, and the push-forward method. The first method is stochastic, whereas the last two are entirely deterministic (don't use random number generators). The PDP Monte-Carlo method combines the numerical simulation of the Markov chain of promoter states with symbolic solutions of the deterministic ODEs and is much faster than the direct Gillespie method [15].

This paper is structured as follows: In Section 2 we recall the justification of piecewise-deterministic approximation using the partial Ω expansion. In Section 3 we introduce a class of piecewise-deterministic models covering gene network. In Section 4.1 we discuss the Monte-Carlo methods for piecewise-deterministic models. In Section 4.2 we discuss the Liouville-master equation approach. In Section 4.3 we discuss the push-forward method. In Section 5 we briefly discuss possible applications of these methods to extracting

information from mRNAs' and proteins' spectra of fluctuations.

2 Partial omega-expansion and the Liouville-master equation

The dynamics of stochastic biochemical networks can be described by a pure jump Markov process [1, 44, 4] (for a general presentation of Markov jump processes see [14]). The state of the network is a vector $X \in \mathbb{N}^n$ whose components represent copy numbers of molecules of various species. Each biochemical reaction is defined by a stoichiometric (jump vector) $\gamma_i \in \mathbb{Z}^n$, $i \in \mathscr{R}$ where $\mathscr{R}$ is the set of biochemical reactions in the model. The occurrence of the reaction i is represented as a jump in the system's state $X \to X + \gamma_i$. Finite, discrete states of gene promoters can be represented by extending the meaning of species to include "places" with finite values of the copy numbers. For instance, on/off gene promoters can be represented by using two places P_{on} and P_{off} with two possible occupancies 0 or 1. The transitions from OFF to ON and back can be represented by a reversible reaction that consumes a particle on the place P_{off} and produces a particle on the place P_{on}.

In the Markov pure jump representation the dynamics of the system is a series of jumps separated by exponentially distributed waiting times [14, 38]. The number of reactions of the type i occurring in the average per unit time is given by the propensity (or rate) function $X \to V_i(X;\mu)$, where μ are kinetic parameters. The time between successive reactions is exponentially distributed with an average $(\sum_{i\in\mathscr{R}} V_i)^{-1}$ and the next reaction is i with probability $V_i/(\sum_{i\in\mathscr{R}} V_i)$. We are interested in the multivariate distribution $p(X,t)$ representing the probability to be in state X at time t. This satisfies the time

dependent master equation:

$$\frac{\partial p}{\partial t}(X,t) = \sum_{i \in \mathscr{R}} [V_i(X - \gamma_i; \mu) p(X - \gamma_i, t) - V_i(X; \mu) p(X, t)]. \tag{2.1}$$

The van Kampen Ω (system size) expansion [51] or equivalently, the central limit theorem [28, 16] lead, in the first order, to deterministic (ODE) approximations and to diffusion (Fokker Planck) approximations, in the second order. The partial omega expansion consists in applying the Ω expansion only to species that are produced in sufficiently large copy numbers [7, 8]. The large copy numbers species denoted X_C are called continuous, because the biochemical reactions change their values gradually, by the accumulation of a large number of small steps. Other species are present only in a few copies per cell (these include the "places" describing promoter states). We denote these species X_D and call them discrete. This leads to a decomposition of the state vector as $X = (X_D, X_C)$ and of all stoichiometric vectors as $\gamma_i = (\gamma_i^D, \gamma_i^C)$, corresponding to discrete and continuous species coordinates. The interactions among discrete and continuous species are suitably described by a partition of the reactions in four sets $\mathscr{R} = \mathscr{R}_D \cup \mathscr{R}_{DC} \cup \mathscr{R}_{CD} \cup \mathscr{R}_C$ [7, 8, 39].

The reactions $\mathscr{R}_D$ act on X_D (the corresponding γ_i have non-zero coordinates on discrete species, $\gamma_i^D \neq 0$) and have propensities depending on X_D only. The reactions $\mathscr{R}_C$ act on X_C (the corresponding γ_i have non-zero coordinates on continuous species, $\gamma_i^C \neq 0$) and have propensities depending on X_C only. The reactions $\mathscr{R}_{DC}$, $\mathscr{R}_{CD}$ act on X_C and X_D, respectively, and their propensities depend on both X_D and X_C.

In this paper we consider gene network models. For each gene, we model the transitions between promoter states, as well as other processes such as transcription, translation, protein folding, and protein

and mRNA degradation. We will consider that the mRNA molecules and proteins are in sufficiently large copy numbers to justify continuous approximations. The only discrete variables are in this case the promoter states. The set $\mathscr{R}_D$ contains transitions between discrete promoter states whose rates do not depend on regulatory proteins. The set $\mathscr{R}_{CD}$ contains transitions between promoter states whose rates depend on concentrations of regulatory proteins. The set $\mathscr{R}_C$ contains translation, maturation (folding), degradation reactions. The set $\mathscr{R}_{DC}$ contains transcription initiation reactions that depend on the promoter state.

We further consider that the copy numbers of continuous species X_C and the propensities of reactions in the sets $\mathscr{R}_C$ and $\mathscr{R}_{DC}$ are "extensive", in other words, scale with the system size Ω, $X_C = \Omega x_c$, $V_i = \Omega v_i(x_c, X_D)$, for $i \in \mathscr{R}_{DC} \cup \mathscr{R}_C$, whereas the propensities of reactions in $\mathscr{R}_D$ and $\mathscr{R}_{CD}$ are considered "intensive" and do not scale with Ω. The Ω dependence is not only a useful mathematical tool, but has also a biological meaning. In proliferating cells, the protein concentrations are important for biochemical reactions and should be maintained by synthesis reactions. This is only possible if the propensities of synthesis reactions (including mRNA synthesis) scale with the size. Rates of monomolecular reactions consuming continuous reactants (for instance, degradation reactions) are proportional to the reactant copy number that scale with the size. Rates of switching reactions between discrete promoter states do no scale with size, unless they are proportional to copy numbers of activator or repressor proteins. For a more complete discussion of these scaling relations we refer to [8].

In rescaled variables, the master equation reads

$$\frac{\partial p}{\partial t}\quad(X_D, x_c, X, t) = \sum_{i\in\mathscr{R}_D\cup\mathscr{R}_{CD}} [V_i(X_D-\gamma_i^D, x_c;\mu)p(X_D-\gamma_i^D, x_c, t) - V_i(X;\mu)p(X_D, x_c, t)] +$$
$$+ \sum_{i\in\mathscr{R}_C\cup\mathscr{R}_{DC}} \Omega[v_i(X_D, x_c-\gamma_i^c/\Omega;\mu)p(X_D, x_c-\gamma_i^C/\Omega, t) - v_i(X, x_c;\mu)p(X_D, x_c, t)]. \tag{2.2}$$

Using the first order Taylor series expansion $v_i(X_D, x_c-\gamma_i^c/\Omega;\mu)p(X_D, x_c-\gamma_i^C/\Omega, t)-v_i(X, x_c;\mu)p(X_D, x_c, t) \approx -\frac{\gamma_i^c}{\Omega}\frac{\partial v_i(X,x_c;\mu)p(X_D,x_c,t)}{\partial x_c}$ we obtain the Liouville-master equation

$$\begin{aligned}\frac{\partial p}{\partial t}(X_D, x_c, t) &= -\frac{\partial[\Phi(X_D, x_c;\mu)p(X_D, x_c, t)]}{\partial x_c} + \sum_{i\in\mathscr{R}_D\cup\mathscr{R}_{CD}} [V_i(X_D-\gamma_i^D, x_c;\mu)p(X_D-\gamma_i^D, x_c, t) - \\ &- V_i(X_D, x_c;\mu)p(X_D, x_c, t)], \text{where} \\ \Phi(X_D, x_c;\mu) &= \sum_{i\in\mathscr{R}_C\cup\mathscr{R}_{DC}} \gamma_i^C v_i(X_D, x_c;\mu).\end{aligned} \tag{2.3}$$

The Liouville-master equation has been used in various fields; in statistical physics of quantum systems [27, 13, 5] or with a different name in statistics and operations research [9]. It describes the time-dependent distribution of a piecewise-deterministic model. Indeed, conditionally on X_D (considering that the discrete variables are fixed) the continuous variables x_c satisfy deterministic, ODEs:

$$\frac{dx_c}{dt} = \Phi(X_D, x_c;\mu). \tag{2.4}$$

The discrete variables follow a pure jump Markov dynamics defined by the stoichiometric vectors γ_i^D and the propensities $V_i(X_D, x_c;\mu)$, $i \in \mathscr{R}_D \cup \mathscr{R}_{CD}$. The noise in the system is produced by the stochastic transitions of the discrete variables. The probability distribution of the continuous variables is advected (transported by the flow defined by the ODEs, see Figure 3) by the deterministic flow $\Phi(X_D, x_c;\mu)$ in the continuous variables space; this explains the advection term $-\frac{\partial}{\partial x_c}[\Phi(X_D, x_c;\mu)p(X_D, x_c, t)]$ in the Liouville-master equation.

3 PDP models of gene circuits

In this section we introduce a family of PDP models that can be used to represent gene networks. They are a special, simplified case of the class of models defined by (2.3). The main simplification is that the propensities of reactions in $\mathscr{R}_{DC}$ depend on X_D and do not depend on x_C.

As an example, we consider a gene circuit with dichotomous noise. This model is made of n_g genes, each one controlled by a promoter with 2 states; ON and OFF. The continuous variables are x_i and y_i; the protein and mRNA level for each gene i, respectively. The discrete variables are two values variables $s_i \in \{0, 1\}$ representing the promoter states (0 stands for OFF and 1 stands for ON). In this model, the set $\mathscr{R}_{DC}$ consists of transcription initiation reactions. Their rates depend on the promoters states (ON of OFF) but do not depend on the protein or mRNA levels. The set $\mathscr{R}_{CD} \cup \mathscr{R}_D$ consists of reversible transitions between the states ON and OFF. The corresponding transition rates are constant when the gene i is constitutive, namely these rates are f_i from ON to OFF and h_i from OFF to ON. When the gene i is activated by the gene j the transition rate from OFF to ON is $f_i x_j$, whereas when the gene i is repressed by the gene j the transition rate from ON to OFF is $h_i x_j$. The discrete variables' dynamics is thus a Markov chain with the state set $\mathcal{M} = \{0, 1\}^{n_g}$. It is convenient to relabel the states from 1 to 2^{n_g} using the lexicographic order on $\mathcal{M}$. For instance, two gene circuits have, in order, the states $1 : (0, 0)$, $2 : (0, 1)$, $3 : (1, 0)$, $4 : (1, 1)$. Also, instead of using reaction propensities, we equivalently provide a transition rate matrix S whose elements are the transition probabilities per time unit between states. For instance, for a two gene circuit where the first gene is constitutive and the second gene is activated by the

first one, we have:

$$S = \begin{pmatrix} 0 & f_1 & f_2 x_1 & 0 \\ h_1 & 0 & 0 & f_2 x_1 \\ h_2 & 0 & 0 & f_1 \\ 0 & h_2 & h_1 & 0 \end{pmatrix}. \tag{3.1}$$

The mRNA and protein variables follow ODE dynamics

$$\begin{aligned} \frac{dx_i}{dt} &= b_i y_i - a_i x_i, \\ \frac{dy_i}{dt} &= k_i(s) - \rho_i y_i, \end{aligned} \tag{3.2}$$

where ρ_i, b_i, a_i, $i \in [1, n_g]$ represent mRNA degradation, translation, and protein degradation rates for the gene i, respectively, and $s = (s_1, s_2, \ldots, s_{n_g})$ is the state of the Markov chain, such that

$$k_i(s) = \begin{cases} k_0 & \text{if } s_i = 0 \\ k_1 & \text{if } s_i = 1 \end{cases}.$$

For the sake of illustration let us consider the simple model of a single constitutive gene controlled by a two state (ON/OFF) promoter. We denote the states of the promoter by 1 and 0 respectively. The transition rate from 0 to 1 is f and from 1 to 0 is h. The protein and mRNA concentrations are x and y, respectively. The transcription initiation rate in the state 1 is k_1 and in the state 0 is $k_0 << k_1$. The translation rate is b. The mRNA and protein degradation rates are ρ and a, respectively. The

master-Liouville equation reads

$$\frac{\partial p}{\partial t}(1,x,y,t) = -\frac{\partial[(by-ax)p(1,x,y,t)]}{\partial x} - \frac{\partial[(k_1-\rho y)p(1,x,y,t)]}{\partial y} + fp(0,x,y,t) - hp(1,x,y,t),$$

$$\frac{\partial p}{\partial t}(0,x,y,t) = -\frac{\partial[(by-ax)p(0,x,y,t)]}{\partial x} - \frac{\partial[(k_0-\rho y)p(0,x,y,t)]}{\partial y} + hp(1,x,y,t) - fp(0,x,y,t). \tag{3.3}$$

The protein and mRNA concentrations follow the ODEs

$$\begin{aligned} \frac{dx}{dt} &= by - ax, \\ \frac{dy}{dt} &= \begin{cases} k_0 - \rho y & \text{if} \quad s = 0 \\ k_1 - \rho y & \text{if} \quad s = 1 \end{cases}. \end{aligned} \tag{3.4}$$

The probability distribution of the promoter state s results from the dynamics of the two state Markov chain

$$\begin{aligned} \frac{dp_0}{dt} &= -fp_0(t) + h(1 - p_0(t)), \\ p_1(t) &= 1 - p_0(t), \end{aligned} \tag{3.5}$$

where $p_0 = \mathbb{P}[s = 0] = \int p(0,x,y)\,dxdy$, $p_1 = \mathbb{P}[s = 1] = \int p(1,x,y)\,dxdy$.

We also define the asymptotic occupancy probabilities $p_0 = \frac{h}{h+f}$ and $p_1 = \frac{f}{h+f}$, representing the probabilities, at steady state, that the promoter state is OFF and ON, respectively.

The single constitutive gene model and the advection fluxes of the Liouville-master equation are illustrated in Figure 3. More complex, two gene circuits models are represented in Figure 4 and their Liouville-master equations are given in the Appendix 1.

4 The numeric methods for solving the Liouville-master equation

In this section we compare several numerical methods for solving the Liouville-master equation in the context of gene networks models. In order to quantify the relative difference between methods we use the L^1 distance between distributions. More precisely, if $p(x)$ and $\tilde{p}(x)$ are probability density functions to be compared, the distance between distributions is

$$d = \int |p(x) - \tilde{p}(x)|\, dx. \tag{4.1}$$

4.1 The PDP Monte-Carlo method

The PDP Monte-Carlo method is based on the direct simulation of the PDP process. A simple algorithm has been proposed in [8]. For the sake of completeness we recall here the main steps of this algorithm.

(1) Set the initial state condition $s = s_0$, $x = x_0$, $y = y_0$, $t = 0$.

(2) Generate a random variable u uniformly distributed in $[0, 1]$,

(3) Integrate the system of differential equations obtained by adding to (3.2) the equation for the survival function F of the waiting time to the next Markov chain transition

$$\begin{cases} \frac{dx_i}{dt} = b_i y_i - a_i x_i, \\ \frac{dy_i}{dt} = k_i(s) - \rho_i y_i, \\ \frac{dF}{dt} = -\lambda(s, x) F, \\ x(0) = x, y(0) = y, F(0) = 1, \end{cases} \tag{4.2}$$

between 0 and τ with the stopping condition

$$F(\tau) = u. \tag{4.3}$$

Here $\lambda(s, x)$ is the sum of transition probabilities in the row corresponding to the state s of the transition rates matrix S.

(4) Generate a second uniform variable v and use it to find the next Markov chain state s_{next}. The decision for the next discrete state is made in the same way as in the Gillespie algorithm [15].

(5) Change the system state $s = s_{next}$, $x = x(\tau)$, $y = y(\tau)$, and the time $t = t + \tau$.

(6) Reiterate the system from (2) with the new state until a time $t_{\max}$ previously defined is reached.

(7) Store samples of x and y at various predefined times.

(8) Reiterate the system from (1) a large number of times. Perform distribution estimates using the resulting statistical ensemble.

A major improvement with respect to [8] can be obtained by using symbolic solutions of (4.2) (see the Appendix 2 for the symbolic solutions) and solving numerically the nonlinear stopping condition (4.3).

4.2 The finite difference (FD) Liouville-master method

The Liouville-master equation is a system of linear, partial differential equations (PDEs) for the probability distributions of the mRNAs and proteins for various states of the gene promoters. The number of PDEs is equal to the number of distinct discrete states of the gene promoters, i.e. it is 2^{n_g}, where n_g is the number of genes in the network.

These equations have to be integrated with boundary conditions in order to control possible mass

accumulation on the integration domain boundaries. The boundary conditions are obtained by setting to zero the advection fluxes pointing towards the boundary. For the one-gene model (3.3), the integration domain is $\mathcal{D} = [\frac{k_0 b}{\rho a}, \frac{k_1 b}{\rho a}] \times [\frac{k_0}{\rho}, \frac{k_1}{\rho}]$ and the boundary conditions read

$$\lim_{y \to k_1/\rho} (k_1 - \rho y) p(1, x, y, t) = 0$$
$$\lim_{y \to k_0/\rho} (k_0 - \rho y) p(0, x, y, t) = 0. \tag{4.4}$$

One can note the limit version of the boundary conditions allow divergence of p_0, p_1 on the boundaries $y = k_0/\rho$ or $y = k_1/\rho$, respectively. For boundaries different than $y = k_0/\rho$ or $y = k_1/\rho$, advection fluxes point towards the interior of the integration domain and one can simply impose zero boundary conditions $p_0 = p_1 = 0$.

As noticed by Marc Kac in a very instructive paper about a piecewise-deterministic random walk [25], in contrast to the Fokker-Planck equations that describe "normal" random-walk diffusion and are parabolic, the Liouville-master equation has hyperbolic nature. General properties of hyperbolic equations, such as finite propagation velocity of perturbations and existence of sharp fronts apply to our equations as well. For our problem, the front discontinuities occur at the domain boundaries and they are handled by the boundary conditions (4.4). Hyperbolicity properties are mainly visible at slow switching and should disappear at fast switching when the Liouville-master equation can be well approximated by a Fokker-Planck equation [39].

In this paper we have used a finite-difference predictor-corrector scheme [47] to compute the solution of the Liouville-master equation. In Figure 5 we compare the distributions for mRNA and proteins resulting from the constitutive gene model (3.3) with the Monte-Carlo simulation of the model. The comparison

is quantitative and uses the distance defined by (4.1). In all our computations, the asymptotic occupancy probability is one half. For slow switching, the mRNA distribution is bimodal, with discontinuities at the modal values k_0/ρ and k_1/ρ values, where the probability density diverges on one side and is zero on the other side. The bimodality resulting from slow switching is well understood and signalled in many other places in the literature (see for instance [45, 23]). We can emphasize that the discontinuity of the solution is a consequence of the hyperbolicity of advection fluxes. A parabolic diffusion flux would not be able to build up such discontinuities and this can be seen in the fast switching distributions that are continuous and unimodal. Interestingly, the protein distributions are unimodal in both cases: slow and fast. A unique discontinuity can be observed at short times in the protein distribution, for a slowly switching promoter.

The bivariate mRNA vs. protein distributions are shown at steady state in Figure 6. The mRNA is positively correlated to the protein as it should be. Interestingly, for slow switching, cells close to half protein maximum expression have strongly bimodal mRNA expression, either vanishing or maximum. This rather unintuitive behaviour results from the difference in lifetime between the mRNA and protein molecules. The mRNA molecule has a short life and can, for slow switching, oscillate between very low and maximum values. The protein is much more stable and integrates these oscillations over a large lifetime. This explains why a cell with half protein maximum copy number can have extremely variable mRNA copy numbers.

4.3 The push-forward method

To introduce the method let us consider the example of a two gene circuit in which the first gene is constitutive and activates the second gene. In other words, the first gene is a transcription factor of the second one.

The push-forward method for gene networks was first introduced in [22]. This method combines Master Equations (MEs) and Random Differential Equations (RDEs). The ME provides the time evolution of the probability distribution of discrete variables. With respect to [22] where the discrete variables were the promoter ON/OFF states and the mRNA copy numbers, here we consider that the discrete variables are solely the promoter's ON/OFF states. The RDEs are differential equations for the mRNAs and proteins in which the promoter states are considered random parameters. The coupling of ME and RDE is another equivalent way to define the piecewise-deterministic process. When the ME is not dependent on the RDE our models are of the same type as those discussed by Mark Kac in relation to the telegrapher's equation [25].

Let us introduce our model beginning with the MEs describing the dynamics of the first switch,

$$\frac{d}{dt}\vec{P}(t) = H_1\vec{P}(t), \tag{4.5}$$

where, $\vec{P}(t) = (P_0(t), P_1(t))^{\mathrm{T}}$ is the probability occupancy vector whose entries are the probabilities to find the first switch in the OFF state ($P_0(t)$) or in the ON state ($P_1(t)$). The infinitesimal stochastic

matrix H_1 is given by:

$$H_1 = \begin{pmatrix} -f_1 & h_1 \\ f_1 & -h_1 \end{pmatrix}. \tag{4.6}$$

This is a basic telegraph process where the rates f_1 and h_1 are time independent and control the reaction OFF (state 0) to ON (state 1) and ON (state 1) to OFF (state 0), respectively. The equation describing mRNA dynamics is a Random Differential Equation (RDE) given by

$$\frac{d}{dt}y_1 = -\rho y_1 + k(s_1(t)), \tag{4.7}$$

where y_1 is a random variable representing the copy number of mRNA in the cell coming from the first gene. The random variable $s_1(t)$ follows the switch statistics, meaning that with probability $P_1(t)$, $s_1(t) = 1$ at time t and $s_1(t) = 0$ with probability $P_0(t)$, again at time t. The production rate of mRNA is a function of the random variable $s_1(t)$ following

$$k(s_1(t)) = \begin{cases} K_1 \text{ if } s_1(t) = 1, \\ K_0 \text{ if } s_1(t) = 0, \end{cases}$$

where K_1 is the highest level of mRNA production and K_0 is the basal one. The third equation describing the activity of the first gene is for the random variable representing the protein density associated to it;

$$\frac{d}{dt}x_1 = -\alpha x_1 + \beta y_1, \tag{4.8}$$

where α is protein degradation rate and β is the translation rate. The last equation for the coupled gene model is the one governing the probability occupancy of the second gene,

$$\frac{d}{dt}\vec{Q}(t) = H_2(t)\vec{Q}(t), \tag{4.9}$$

where $\vec{Q}(t) = (Q_0(t), Q_1(t))^{\mathrm{T}}$ encodes, in its entries, the information about the probability to find the second gene ON ($Q_1(t)$) or OFF ($Q_0(t)$). The matrix $H_2(t)$ is given by

$$H_2(t) = \begin{pmatrix} -f_2 x_1(t) & h_2 \\ f_2 x_1(t) & -h_2 \end{pmatrix}. \tag{4.10}$$

In the model at hand, the main source of stochasticity is the switching ON and OFF of the gene. This noise is transmitted to mRNA synthesis process through the rate $k(s_1(t))$ which is a function of a random variable ($s_1(t)$) and, so, a random variable itself. The first step of the push-forward method is to compute the time dependent distribution probability of mRNA molecules $y_1(t)$ (which is perturbed by the random variable $s_1(t)$) once the probability distribution of the perturbation is known. To do so, we begin by presenting the solutions of Eqs (4.5),

$$\vec{P}(t) = \exp[(UDU^{-1})(t - t_0)]\vec{P}(t_0) = Ue^{(D(t-t_0))}U^{-1}\vec{P}(t_0), \tag{4.11}$$

where $\vec{P}(t_0)$ encodes the initial configuration (given at $t = t_0$) of the switch, and the matrices are

$$U = \begin{pmatrix} -1 & \frac{h}{f+h} \\ 1 & \frac{f}{f+h} \end{pmatrix}, \; U^{-1} = \begin{pmatrix} -\frac{f}{f+h} & \frac{h}{f+h} \\ 1 & 1 \end{pmatrix} \text{, and } D = \begin{pmatrix} -f-h & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.12}$$

Explicitly, the solutions are given as

$$\begin{aligned} P_0(t) &= p_0 + (P_0(t_0) - p_0)e^{-(h+f)(t-t_0)}, \\ P_1(t) &= p_1 + (P_1(t_0) - p_1)e^{-(h+f)(t-t_0)}, \end{aligned} \tag{4.13}$$

where $p_0 = h/(f+h)$ and $p_1 = f/(f+h)$ are the asymptotic occupancy probabilities to find the gene OFF

and ON, respectively. Going on, we present the formal solution of the RDE governing mRNA dynamics,

$$y_1(\tau) = y_1(\tau_0)e^{-(\tau-\tau_0)} + \int_{\tau_0}^{\tau} d\tau' e^{-(\tau-\tau')}[k_0(1 - s_1(\tau')) + k_1 s_1(\tau')], \tag{4.14}$$

where we have rescaled time t by the mRNA degradation rate and introduced the new time parameter, $\tau = t\rho$, and also the dimensionless parameters $k_0 = K_0/\rho$ and $k_1 = K_1/\rho$. Note that the integral (4.14) is a basic Riemann integral, such that, if we consider a sufficiently fine partition $[\tau_0, \tau_1, ..., \tau_{N-1}, \tau_N]$ of the integral interval $[\tau_0, \tau]$, where $\tau_N = \tau$, we can assume that $s_1(\tau')$ is constant inside each specific partition $[\tau_j, \tau_{j+1}]$ and the integral in (4.14) is approximated by

$$\begin{aligned}\int_{\tau_0}^{\tau} d\tau' e^{-(\tau-\tau')}[k_0(1 - s_1(\tau')) + k_1 s_1(\tau')] &\approx \sum_{j=0}^{N-1} [k_0(1 - s_1(\tau_j)) + k_1 s_1(\tau_j)] \int_{\tau_j}^{\tau_{j+1}} d\tau' e^{-(\tau-\tau')} \\ &= e^{-\tau} \sum_{j=0}^{N-1} [(k_0(1 - s_1(\tau_j)) + k_1 s_1(\tau_j))(e^{\tau_{j+1}} - e^{\tau_j})].\end{aligned} \tag{4.15}$$

It is worthwhile to remember that $s_1(\tau)$ is a jumping process assuming, at each instant of time, just one of the values, 0 or 1, with probability $P_0(\tau)$ or $P_1(\tau)$, respectively. Now the solution of $y_1(\tau)$ is given as a function of sequences of $s_1(\tau_j)$, with $j = 0, ..., N-1$. These sequences are strings of zeros and ones and we must consider all of them. For instance, if the number of partitions is N, we will have $2^{(N-1)}$ possible sequences and each one will lead to a different value of $y_1(\tau)$. The remaining task is to assign the correct weight for each one of these values of $y_1(\tau)$. This is achieved by using the occurrence probability of each possible string of zeros and ones defining the values of $y_1(\tau)$. The occurrence probability is given by the joint distribution probability of having $s_1(\tau_j) = 0$ or 1 at time τ_j for $j = 0, ..., N-1$ which is

$$\Pr(s_1(\tau_0), \tau_0; ...; s_1(\tau_{N-1}), \tau_{N-1}) = \mathrm{Prc}(s_1(\tau_{N-1}), \tau_{N-1}|s_1(\tau_{N-2}), \tau_{N-2})...\mathrm{Prc}(s_1(\tau_1), \tau_1|s_1(\tau_0), \tau_0)P_{s_1(\tau_0)}(\tau_0), \tag{4.16}$$

where $P_{s_1(\tau_0)}(\tau_0)$ is the occupancy probability to find the gene in state $s_1(\tau_0)$ at time τ_0, as per (4.13), and $\mathrm{Prc}(\ldots|\ldots)$ stands for the conditional probability for the telegraph process. For instance,

$$\mathrm{Prc}(s_1(\tau_{N-1}), \tau_{N-1}|s_1(\tau_{N-2}), \tau_{N-2}),$$

encodes the transition probability from state $s_1(\tau_{N-2})$ to state $s_1(\tau_{N-1})$ during the time interval $\tau_{N-1} - \tau_{N-2}$ knowing that at τ_{N-2} the system was with probability one in state $s_1(\tau_{N-2})$. The conditional probabilities for the telegraph process, which is a Markovian process, are easily obtained by chosen the appropriate initial configuration:

$$\begin{aligned}
\mathrm{Prc}(0,\tau|0,\tau_0) &= p_0 + (1-p_0)e^{-\epsilon(\tau-\tau_0)}, \\
\mathrm{Prc}(0,\tau|1,\tau_0) &= p_0 - p_0 e^{-\epsilon(\tau-\tau_0)}, \\
\mathrm{Prc}(1,\tau|0,\tau_0) &= p_1 - p_1 e^{-\epsilon(\tau-\tau_0)}, \\
\mathrm{Prc}(1,\tau|1,\tau_0) &= p_1 + (1-p_1)e^{-\epsilon(\tau-\tau_0)},
\end{aligned} \tag{4.17}$$

where the new parameter $\epsilon = (f+h)/\rho$ measures the flexibility of the switch. With the conditional probabilities and (4.15) at hand, we can calculate the time dependent probability distribution for the mRNA density. We have considered two examples: one is the fast switch regime ($\epsilon > 1$) and, the other, slow switching ($\epsilon < 1$).

Protein distribution in time can be obtained in the same fashion as the one for mRNA. The general solution for protein density is:

$$\begin{aligned}
x_1(\tau) = x_1(\tau_0)e^{-a(\tau-\tau_0)} &+ \frac{b\,y_1(\tau_0)e^{-\tau_0}}{a-1}\left[e^{-(\tau-\tau_0)} - e^{-a(\tau-\tau_0)}\right] \\
&+ b\int_{\tau_0}^{\tau} d\tau' e^{-a(\tau-\tau')}\int_0^{\tau'} d\tau'' e^{-(\tau'-\tau'')}[k_0(1-s_1(\tau'')) + k_1 s_1(\tau'')],
\end{aligned} \tag{4.18}$$

where the rescaled parameters are given by $a = \alpha/\rho$ and $b = \beta/\rho$. The push-forward method can be applied to obtain the time dependent distribution probability for protein density, in an analogous way as for mRNA. The integral that must be partitioned is that over τ', in the interval $[\tau_0, \tau]$

$$\begin{aligned} be^{-a\tau}\int_{\tau_0}^{\tau} d\tau' e^{(a-1)\tau'}\int_0^{\tau'} d\tau'' e^{\tau''}k(s_1(\tau'')) &\approx be^{-a\tau}\sum_{j=0}^{N-1} k(s_1(\tau_j))\int_{\tau_j}^{\tau_{j+1}} d\tau' e^{(a-1)\tau'}\int_0^{\tau'} d\tau'' e^{\tau''} \\ &= be^{-a\tau}\sum_{j=0}^{N-1} k(s_1(\tau_j))\left[\frac{e^{a\tau_{j+1}}-e^{a\tau_j}}{a} - \frac{e^{(a-1)\tau_{j+1}}-e^{(a-1)\tau_j}}{a-1}\right], \end{aligned} \tag{4.19}$$

where, we have used the definition $k(s_1(\tau)) = k_0(1 - s_1(\tau)) + k_1 s_1(\tau)$ to simplify the notation. As before, we have illustrated our method by calculating the protein density for the same two regimes of switch flexibility.

To analyze the influence of the first gene on the second one we have assumed that the action of the first gene is to activate the second (see (4.10)). To do so, instead of solving the RDE describing the activity of the second gene (it is an RDE because the perturbation $x_1(\tau)$ is a random variable) we have analyzed the mean value of the occupancy probability of the second gene whose dynamics is given by

$$\frac{d}{d\tau}\langle Q_0(\tau)\rangle = -(h_2 + f_2\langle x_1(\tau)\rangle)\langle Q_0(\tau)\rangle + h_2, \tag{4.20}$$

and $\langle Q_1(\tau)\rangle = 1 - \langle Q_0(\tau)\rangle$. The general solution for $\langle Q_0(\tau)\rangle$ is given by

$$\langle Q_0(\tau)\rangle = \langle Q_0(0)\rangle e^{-\int_0^{\tau} d\tau'(h_2+f_2\langle x_1(\tau')\rangle)} + \int_0^{\tau} d\tau' h_2 e^{\int_0^{\tau'} d\tau''(h_2+f_2\langle x_1(\tau'')\rangle)}. \tag{4.21}$$

In Appendix 3 we show, in detail, how to obtain the exact functional shape of $\langle x_1(\tau)\rangle$. Nevertheless, its structure is $\langle x_1(\tau)\rangle = r_0 + r_1 e^{-\tau} + r_2 e^{-a\tau} + r_3 e^{-\epsilon\tau}$ and, because of this, the integral in (4.21) cannot be evaluated analytically and a numerical evaluation must be performed. This will also be the case for the

conditional probabilities that will be expressed as

$$\mathrm{Prc}(s_2(\tau_j),\tau_j|s_2(\tau_{j-1}),\tau_{j-1}) = s_2(\tau_{j-1})e^{-\int_{\tau_{j-1}}^{\tau_j} d\tau'(h_2+f_2\langle x_1(\tau')\rangle)} + \int_{\tau_{j-1}}^{\tau_j} d\tau' h_2 e^{\int_0^{\tau'} d\tau''(h_2+f_2\langle x_1(\tau'')\rangle)}, \quad (4.22)$$

where we have set $\langle Q_i(\tau_{j-1})\rangle = s_2(\tau_{j-1})$ (with $i = 0$ or 1) expressing the fact that at the instant of time τ_{j-1} the gene 2 is in state i with probability $s_2(\tau_{j-1})$ (which is 0 or 1), as we have done for the gene 1. The difference is that for gene 1 we have closed solutions for the occupancy probabilities and the conditional probabilities were derived analytically from these solutions in (4.17). However, here, as said before, we must perform the integrals in (4.22) numerically.

With (4.21) and (4.22) at hand we are in position to obtain the distributions for mRNA and proteins associated with gene 2 in the same fashion as we did for the gene 1. The only difference is that the mRNA and protein copy numbers will be obtained using (4.15) and (4.19) just by changing the index 1 for 2 ($s_1(\tau) \to s_2(\tau)$, $y_1(\tau) \to y_2(\tau)$, and $x_1(\tau) \to x_2(\tau)$). A last comment regards the parameter space: as we did for the gene 1, we have redefined the parameter space of gene 2 and introduced the more biological parameters; the asymptotic occupancy probabilities (q_0, q_1) and the flexibility parameter (σ). The new parameters are expressed in terms of the old ones as:

$$q_0 = \frac{h_2}{f_2 r_0 + h_2}, \quad q_1 = \frac{f_2 r_0}{f_2 r_0 + h_2}, \quad \sigma = \frac{f_2 r_0 + h_2}{\rho}. \quad (4.23)$$

In figure 7 we compare the distributions for mRNA and proteins associated with gene 2 with the direct simulation of the model. We have done that in four distinct situations when the first and the second genes are fast and slow. For all the cases the first and second gene are OFF at $\tau = 0$ and both have asymptotic

occupancy probabilities equal to 1/2. The comparison is quantitative and uses the distance defined by (4.1).

5 Testing the propensity of mRNA and protein intrinsic fluctuations to reveal gene/gene interactions.

In this paper we used mathematical models to predict mRNA and protein "intrinsic" fluctuations (by "intrinsic" we understand fluctuations generated by the stochastic gene network dynamics). An important question in this context is if "intrinsic" gene expression fluctuations can be used to reverse engineer gene networks. "Extrinsic" fluctuations (by "extrinsic" we understand perturbations of gene networks coming from their environment) of the transcriptome were extensively used in the past to reconstruct gene networks using correlation or mutual information (for a popular method see [31]). A few reverse-engineering studies obtained gene regulation parameters from intrinsic gene expression fluctuations [12, 34]. Quantifying intrinsic fluctuations ask for single cell measurements of several genes. A variety of technologies are now ready to take this challenge: single cell sequencing [46], single-cell RNA microscopy [48], various versions of time-lapse microscopy [33], fluorescence correlation microscopy [12]. It is therefore important to test the propensity of expression fluctuations to discriminate between various gene network architectures.

The predictions of our PDP models are provided as multivariate distributions of mRNAs and proteins copy numbers for one or several genes. These predictions can be directly compared with results obtained with molecular biology and biophysics experimental methods.

First, one would like to test if the differences between distributions predicted with various gene circuit models are significant and therefore can be used to discriminate between gene circuit models. To this end, we generated bivariate proteins and mRNA distributions for five different gene circuits like in Figure 4. The visual inspection of Figure 8, suggests that mRNA distributions resulting from five different gene circuits, with gene/gene interactions that differ by their signs, are very similar in the same regime of switching (fast or slow) of the genes. The mRNA distributions discriminate between model parameters (fast or slow swithcing) but do not discriminate between circuit architectures. The protein bivariate distributions are shown in Figure 9. They differ strongly from mRNA distributions and discriminate between both parameters and architectures.

As visual colormap differences may be judged subjective, we developed a quantitative test for the discriminant power. This test is based on the distance d defined by (4.1). We have computed d pairwise, for mRNA and for protein distributions produced from different gene circuits at steady state. In order to test if these distances are significantly large we compared them to distribution of distances between random samples generated by the same gene circuit model for a fixed number of cells. The result of this comparison is shown in Figure 10 for the two gene circuits $G_1 \to G_2$ and $G_1 \dashv G_2$ that differ by the sign of the interaction; one can notice that the protein fluctuation based distance is significant, whereas the mRNA fluctuation distance is not, both for slow/slow and fast/fast genes.

We have also tested the significance of the correlation computed from bivariate mRNA or protein distributions. A simple gene reconstruction method is to consider that genes interact if the correlation coefficient is significantly different from zero. We have computed the Bravais-Pearson correlation coef-

ficient from bivariate mRNA and protein samples generated with our PDP model, at steady state and for increasing numbers of cells N_c. For both models $G_1 \rightarrow G_2$ and $G_1 \dashv G_2$ a significant (upper tail probability $p < 5\%$) protein/protein correlation is obtained for moderate cell populations ($N_c > 100$ for $p < 5\%$, see Figure 10). In order to obtain significant mRNA/mRNA correlation one has to use very large numbers of cells ($N_c > 1000$ for $p < 5\%$, see Figure 10). This is possible for single cell sequencing and flow cytometry (with the drawback of the lack of precision in estimating the mRNA copy numbers) but is very difficult for techniques such as MS2 tagging microscopy, or time-lapse microscopy.

6 Discussion

We have discussed three methods to compute time-dependent distributions of mRNA and protein copy numbers generated by gene networks. All the three methods are much faster than the Gillespie exact chemical master equation. The finite difference Liouville-master equation method is precise and fast for small models. Simple (non-adaptive) finite difference schemes, however, are demanding in terms of space and time resolution leading to computer memory limitations. In future implementations of the Liouville-master equation method we plan to use spectral methods for bypassing these limitations. The push-forward method is not as precise as the Liouville-master equation, but it is much more stable, even at low resolution. The differences between execution times of various methods are illustrated in the Table1 indicating that the push-forward method is the fastest one. The performance of the push-forward method results from the reduced cardinality of the discrete phase space (2 states for one gene, 4 states for 2 genes) which is an improvement with respect to the initial version in [22]. Further improvements, lifting the

mean-field approximation for gene coupling will be presented elsewhere.

Piecewise-deterministic models are valuable tools for understanding stochastic gene expression in a wide spectrum of regimes, covering both slow and fast switching. The source of stochasticity in such models are the random transitions between discrete promoter states; a phenomenon usually associated with transcriptional bursting. In this paper we have only discussed dichotomous noise (ON/OFF promoters), however, as seen in Section 2, our methods work also for promoters with more than two states.

As application of our numerical methods we tested the capacity of mRNA and protein copy numbers fluctuations to unravel differences between gene circuits architectures. We showed that protein fluctuations are sensitive to differences of architecture and that protein-protein correlation reveals gene-gene interactions for moderate cell population sizes (100 cells). In contrast, mRNA fluctuations are much less sensible to differences in circuit architecture and mRNA-mRNA correlation is small, even for interacting genes. This reinforces the already well established conclusion that proteome contains much more information than the transcriptome. In the past we used the spectrum of protein copy number fluctuations to extract information about promoter repression mechanisms [12]. The difference in behavior of the mRNA and protein fluctuations can be explained by the fact that the mRNA half life is usually much shorter than the protein half life. Gene-gene interactions are mediated by proteins that slowly modulate the gene switching times. Proteins follow these slow modulations, which results in significant protein-protein correlation. mRNA dynamics is permanently submitted to the faster (uncorrelated) gene switching, which explains the lower correlation of steady state mRNA fluctuations. This suggests that reconstruction of gene networks from mRNA intrinsic fluctuations is submitted to severe limitations. More general results

concerning these limitations will be presented elsewhere.

Appendix1: Two gene circuits

The Liouville-master equations for two gene circuits are systems of four PDEs in four variables x_1, y_1, x_2, y_2 representing the protein and mRNA of the first gene G_1 and second gene G_2, respectively. There are five different types of such circuits $G_1 \to G_2$, $G_1 \dashv G_2$, $G_1 \to G_2 \to G_1$, $G_1 \to G_2 \dashv G_1$, $G_1 \dashv G_2 \dashv G_1$ where $\to$, $\dashv$ stand for activation and repression, respectively.

The set of two promoters have four discrete states $1 = (0,0)$, $2 = (1,0)$, $3 = (0,1)$, $4 = (1,1)$. The respective probabilities densities are $p_1(x_1, y_1, x_2, y_2, t)$, $p_2(x_1, y_1, x_2, y_2, t)$, $p_3(x_1, y_1, x_2, y_2, t)$, $p_4(x_1, y_1, x_2, y_2, t)$.

The corresponding Liouville-master equations are the following:

For the circuit $G_1 \to G_2$.

$$\begin{aligned}
\frac{\partial p_1}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_1]}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_1]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_1]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_1]}{\partial y_2} + h_2 p_3 + h_1 p_2 - (f_1 + f_2 x_1)p_1, \\
\frac{\partial p_2}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_2]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_2]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_2]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_2]}{\partial y_2} + f_1 p_1 + h_2 p_4 - (h_1 + f_2 x_1)p_2, \\
\frac{\partial p_3}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_3}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_3]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_3]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_3]}{\partial y_2} + h_1 p_4 + f_2 x_1 p_1 - (h_2 + f_1)p_3, \\
\frac{\partial p_4}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_4]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_4]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_4]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_4]}{\partial y_2} + f_1 p_3 + f_2 x_1 p_2 - (h_1 + h_2)p_4.
\end{aligned}$$

For the circuit $G_1 \dashv G_2$.

$$\begin{aligned}
\frac{\partial p_1}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_1]}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_1]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_1]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_1]}{\partial y_2} + h_2 x_1 p_3 + h_1 p_2 - (f_1 + f_2)p_1, \\
\frac{\partial p_2}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_2]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_2]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_2]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_2]}{\partial y_2} + f_1 p_1 + h_2 x_1 p_4 - (h_1 + f_2)p_2, \\
\frac{\partial p_3}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_3}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_3]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_3]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_3]}{\partial y_2} + h_1 p_4 + f_2 p_1 - (h_2 x_1 + f_1)p_3, \\
\frac{\partial p_4}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_4]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_4]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_4]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_4]}{\partial y_2} + f_1 p_3 + f_2 p_2 - (h_1 + h_2 x_1)p_4.
\end{aligned}$$

For the circuit $G_1 \to G_2 \to G_1$.

$$
\begin{aligned}
\frac{\partial p_1}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_1]}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_1]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_1]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_1]}{\partial y_2} + h_2 p_3 + h_1 p_2 - (f_1 x_2 + f_2 x_1)p_1, \\
\frac{\partial p_2}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_2]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_2]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_2]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_2]}{\partial y_2} + f_1 x_2 p_1 + h_2 p_4 - (h_1 + f_2 x_1)p_2, \\
\frac{\partial p_3}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_3}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_3]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_3]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_3]}{\partial y_2} + h_1 p_4 + f_2 x_1 p_1 - (h_2 + f_1 x_2)p_3, \\
\frac{\partial p_4}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_4]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_4]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_4]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_4]}{\partial y_2} + f_1 x_2 p_3 + f_2 x_1 p_2 - (h_1 + h_2)p_4.
\end{aligned}
$$

For the circuit $G_1 \to G_2 \dashv G_1$.

$$
\begin{aligned}
\frac{\partial p_1}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_1]}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_1]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_1]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_1]}{\partial y_2} + h_2 p_3 + h_1 x_2 p_2 - (f_1 + f_2 x_1)p_1, \\
\frac{\partial p_2}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_2]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_2]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_2]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_2]}{\partial y_2} + f_1 p_1 + h_2 p_4 - (h_1 x_2 + f_2 x_1)p_2, \\
\frac{\partial p_3}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_3}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_3]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_3]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_3]}{\partial y_2} + h_1 x_2 p_4 + f_2 x_1 p_1 - (h_2 + f_1)p_3, \\
\frac{\partial p_4}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_4]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_4]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_4]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_4]}{\partial y_2} + f_1 p_3 + f_2 x_1 p_2 - (h_1 x_2 + h_2)p_4.
\end{aligned}
$$

For the circuit $G_1 \dashv G_2 \dashv G_1$.

$$
\begin{aligned}
\frac{\partial p_1}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_1]}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_1]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_1]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_1]}{\partial y_2} + h_2 x_1 p_3 + h_1 x_2 p_2 - (f_1 + f_2)p_1, \\
\frac{\partial p_2}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_2]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_2]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_2]}{\partial x_2} - \frac{\partial[(k_0 - \rho y_2)p_2]}{\partial y_2} + f_1 p_1 + h_2 x_1 p_4 - (h_1 x_2 + f_2)p_2, \\
\frac{\partial p_3}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_3}{\partial x_1} - \frac{\partial[(k_0 - \rho y_1)p_3]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_3]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_3]}{\partial y_2} + h_1 x_2 p_4 + f_2 p_1 - (h_2 x_1 + f_1)p_3, \\
\frac{\partial p_4}{\partial t} &= -\frac{\partial[(by_1 - ax_1)p_4]}{\partial x_1} - \frac{\partial[(k_1 - \rho y_1)p_4]}{\partial y_1} - \frac{\partial[(by_2 - ax_2)p_4]}{\partial x_2} - \frac{\partial[(k_1 - \rho y_2)p_4]}{\partial y_2} + f_1 p_3 + f_2 p_2 - (h_1 x_2 + h_2 x_1)p_4.
\end{aligned}
$$

Appendix2: Analytic solutions for one step deterministic propagation.

For gene circuits (4.2) read

$$\frac{dx_i}{dt} = b_i y_i - a_i x_i,$$
$$\frac{dy_i}{dt} = k_i(s) - \rho_i y_i,$$
$$\frac{dF}{dt} = -[c_0(s) + \sum_{i=1}^{n_g} c_i(s) x_i]F, \tag{6.1}$$

where $c_i(s)$, $i \in [0, n_g]$ are positive constants depending on the discrete Markov chain state s.

The solution of this system is straightforward

$$y_i(t) = \tfrac{k_i(s)}{\rho_i} + [y_i(0) - \tfrac{k_i(s)}{\rho_i}] \exp(-\rho_i t),$$

$$x_i(t) = d_0(i) + d_1(i) \exp(-\rho_i t) + d_2(i) \exp(-a_i t),$$

$$\log F(t) = -(c_0(s) + \textstyle\sum_{i=1}^{n_g} c_i(s) d_0(i))t - \tfrac{1-\exp(-\rho_i t)}{\rho_i}[\sum_{i=1}^{n_g} c_i(s) d_1(i)] - \tfrac{1-\exp(-a_i t)}{a_i}[\sum_{i=1}^{n_g} c_i(s) d_2(i)], \tag{6.2}$$

where $d_0(i) = \frac{b_i k_i(s)}{a_i \rho_i}$, $d_1(i) = \frac{b_i[y_i(0) - \frac{k_i(s)}{\rho_i}]}{a_i - \rho_i}$, $d_2(i) = x_i(0) - d_0(i) - d_1(i)$, $\tau = t - t_j$.

For the single constitutive gene model $c_0(0) = f$, $c_0(1) = h$, $c_1(0) = c_1(1) = 0$, $F(\tau) = \exp[-(f+h)\tau]$. The constants c_i, $i \in \{0, 1, 2\}$ for the two gene models are given in the Table 2. For a general gene network, the constants table is precalculated symbolically and used to generate automatically the simulation code.

Appendix3: Supplementary information for the push-forward method.

In order to obtain an expression for $\langle x_1(\tau)\rangle$ we must solve a set of ordinary differential equations:

$$\begin{aligned} \frac{d}{d\tau}\langle y_1(\tau)\rangle &= -\langle y_1(\tau)\rangle + \langle k(s_1(\tau))\rangle, \\ \frac{d}{d\tau}\langle x_1(\tau)\rangle &= -a\langle y_1(\tau)\rangle + b\langle x_1(\tau)\rangle, \end{aligned} \tag{6.3}$$

where $\langle k(s_1(\tau))\rangle = k_0 P_0(\tau) + k_1 P_1(\tau)$. The solution for $\langle y_1(\tau)\rangle$ is:

$$\begin{aligned} \langle y_1(\tau)\rangle &= k_0 p_0 + k_1 p_1 + \left(n_0 + \frac{k_0[P_0(0) - \epsilon p_0]}{\epsilon - 1} + \frac{k_1[P_1(0) - \epsilon p_1]}{\epsilon - 1}\right) e^{-\tau} \\ &+ \left(\frac{k_0[p_0 - P_0(0)]}{\epsilon - 1} + \frac{k_1[p_1 - P_1(0)]}{\epsilon - 1}\right) e^{-\epsilon\tau}. \end{aligned} \tag{6.4}$$

The general solution for protein mean value has the structure, $\langle x_1(\tau)\rangle = r_0 + r_1 e^{-\tau} + r_2 e^{-a\tau} + r_3 e^{-\epsilon\tau}$, where the coefficients are given by :

$$\begin{aligned} r_0 &= \frac{b}{a}(k_0 p_0 + k_1 p_1), \\ r_1 &= \frac{b}{a-1}\left(n_0 + \frac{k_0(P_0(0) - \epsilon p_0)}{\epsilon - 1} + \frac{k_1(P_1(0) - \epsilon p_1)}{\epsilon - 1}\right), \\ r_2 &= \frac{b}{a-1}\left(\frac{m_0(a-1)}{b} + n_0 + \frac{k_0(aP_0(0) - \epsilon p_0)}{a(a-\epsilon)} + \frac{k_1(aP_1(0) - \epsilon p_1)}{a(a-\epsilon)}\right), \\ r_4 &= \frac{b(k_0(p_0 - P_0(0)) + k_1(p_1 - P_1(0)))}{(\epsilon - 1)(a - \epsilon)}. \end{aligned} \tag{6.5}$$

References

[1] M. Delbrück . Statistical fluctuations in autocatalytic reactions. *J.Chem.Phys.*, 8:120–124, 1940.

[2] Bruce Alberts. *Molecular biology of the cell.* Garland science, 2017.

[3] Nathalie Q Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, 2004.

[4] A.F. Bartholomay. *A Stochastic Approach to Chemical Reaction Kinetics.* Phd thesis, Harvard University, 1957.

[5] Heinz-Peter Breuer and Francesco Petruccione. Stochastic dynamics of quantum jumps. *Physical Review E*, 52(1):428, 1995.

[6] Long Cai, Nir Friedman, and X Sunney Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440(7082):358, 2006.

[7] A. Crudu, A. Debussche, A. Muller, and O. Radulescu. Convergence of stochastic gene networks to hybrid piecewise deterministic processes. *Annals of Applied Probability*, 22:1822–1859, 2012.

[8] A. Crudu, A. Debussche, and O. Radulescu. Hybrid stochastic simplifications for multiscale gene networks. *BMC Systems Biology*, 3(1):89, 2009.

[9] M.H.A. Davis. *Markov Models and Optimization.* Chapman and Hall, London, 1993.

[10] Michael Dean, Tito Fojo, and Susan Bates. Tumour stem cells and drug resistance. *Nature Reviews Cancer*, 5(4):275, 2005.

[11] Michael B Elowitz, Arnold J Levine, Eric D Siggia, and Peter S Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[12] Le Coq D. Jules M. Aymerich S. Radulescu O. Declerck N. Ferguson, M.L. and C.A. Royer. Reconciling molecular regulatory mechanisms with noise patterns of bacterial metabolic promoters in induced and repressed states. *Proceedings of the National Academy of Sciences USA*, 109:155, 2012.

[13] Ralph Gebauer and Roberto Car. Kinetic theory of quantum transport at the nanoscale. *Physical Review B*, 70(12):125324, 2004.

[14] I.I. Gikhman and A.V. Skorokhod. *Introduction to the Theory of Random Processes.* W.B.Saunders Co., Philadelphia, 1969.

[15] Daniel T. Gillespie. *J.Comput.Phys.*, 22:403, 1976.

[16] Daniel T. Gillespie. The chemical langevin equation. *J.Chem.Phys*, 113:297–306, 2000.

[17] Eric L. Haseltine and James B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J.Chem.Phys.*, 117:6959–6969, 2002.

[18] Charlotte Hirschmann-Jax, Aaron E Foster, Gerald G Wulf, Jed G Nuchtern, Thomas W Jax, Ulrich Gobel, Margaret A Goodell, and Malcolm K Brenner. A distinct "side population" of cells with high drug efflux capacity in human tumor cells. *Proceedings of the National Academy of Sciences of the United States of America*, 101(39):14228–14233, 2004.

[19] Caitriona Holohan, Sandra Van Schaeybroeck, Daniel B Longley, and Patrick G Johnston. Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer*, 13(10):714, 2013.

[20] José E. M. Hornos, Daniel Schultz, Guilherme C. P. Innocentini, J. Wang, Aleksandra M. Walczak, José N. Onuchic, and Peter G. Wolynes. Self-regulating gene: An exact solution. *Phys. Rev. E*, 72:051907, 2005.

[21] Guilherme C. P. Innocentini and José E. M. Hornos. Modeling stochastic gene expression under repression. *J. Math. Biol.*, 55:413–431, 2007.

[22] Guilherme CP Innocentini, Michael Forger, Ovidiu Radulescu, and Fernando Antoneli. Protein synthesis driven by dynamical stochastic transcription. *Bulletin of mathematical biology*, 78(1):110–131, 2016.

[23] Guilherme da Costa Pereira Innocentini, Michael Forger, Alexandre Ferreira Ramos, Ovidiu Radulescu, and José Eduardo Martinho Hornos. Multimodality and flexibility of stochastic gene expression. *Bulletin of mathematical biology*, 75(12):2600–2630, 2013.

[24] François Jacob and Jacques Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology*, 3(3):318–356, 1961.

[25] Mark Kac. A stochastic model related to the telegrapher's equation. *The Rocky Mountain Journal of Mathematics*, 4(3):497–509, 1974.

[26] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467, 1969.

[27] Ryogo Kubo. Stochastic Liouville equations. *Journal of Mathematical Physics*, 4(2):174–183, 1963.

[28] Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump markov process. *J.Appl.Prob.*, 7:49–58, 1970.

[29] Thomas G. Kurtz. Limit theorems for sequences of jump markov processes approximating ordinary differential processes. *J.Appl.Prob.*, 8:344–356, 1971.

[30] Daniel R Larson, Robert H Singer, and Daniel Zenklusen. A single molecule view of gene expression. *Trends in cell biology*, 19(11):630–637, 2009.

[31] Adam A Margolin, Ilya Nemenman, Katia Basso, Chris Wiggins, Gustavo Stolovitzky, Riccardo Dalla Favera, and Andrea Califano. Aracne: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. In *BMC bioinformatics*, volume 7, page S7. BioMed Central, 2006.

[32] Harley H McAdams and Adam Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.

[33] Alice Moussy, Jérémie Cosette, Romuald Parmentier, Cindy da Silva, Guillaume Corre, Angelique Richard, Olivier Gandrillon, Daniel Stockholm, and András Páldi. Integrated time-lapse and single-cell transcription studies highlight the variable and dynamic nature of human hematopoietic cell fate commitment. *PLoS biology*, 15(7):e2001867, 2017.

[34] Brian Munsky, Gregor Neuert, and Alexander Van Oudenaarden. Using gene expression noise to understand gene regulation. *Science*, 336(6078):183–187, 2012.

[35] Vivek K Mutalik, Joao C Guimaraes, Guillaume Cambray, Colin Lam, Marc Juul Christoffersen, Quynh-Anh Mai, Andrew B Tran, Morgan Paull, Jay D Keasling, Adam P Arkin, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nature methods*, 10(4):354, 2013.

[36] Damien Nicolas, Nick E Phillips, and Felix Naef. What shapes eukaryotic transcriptional bursting? *Molecular BioSystems*, 13(7):1280–1290, 2017.

[37] Jean Peccoud and Bernard Ycart. Markovian modeling of gene-product synthesis. *Theoretical population biology*, 48(2):222–234, 1995.

[38] O. Radulescu, A. Muller, and A. Crudu. Théorèmes limites pour des processus de markov à sauts. synthèse des resultats et applications en biologie moleculaire. *Technique et Science Informatique*, 26:443–469, 2007.

[39] Ovidiu Radulescu, Guilherme CP Innocentini, and José Eduardo M Hornos. Relating network rigidity, time scale hierarchies, and expression noise in gene networks. *Physical Review E*, 85(4):041919, 2012.

[40] Arjun Raj, Charles S Peskin, Daniel Tranchina, Diana Y Vargas, and Sanjay Tyagi. Stochastic mrna synthesis in mammalian cells. *PLoS biology*, 4(10):e309, 2006.

[41] Arjun Raj and Alexander van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135(2):216–226, 2008.

[42] Alexandre Ferreira Ramos, GCP Innocentini, and José Eduardo Martinho Hornos. Exact time-dependent solutions for a self-regulating gene. *Physical Review E*, 83(6):062902, 2011.

[43] Brandon S Razooky, Anand Pai, Katherine Aull, Igor M Rouzine, and Leor S Weinberger. A hardwired hiv latency program. *Cell*, 160(5):990–1001, 2015.

[44] A. Rényi. Betrachtung chemischer reaktionen mit hilfe der theorie der stochastichen prozesse. *Magyar Tud.Akad.Alkalm.Mat.Int.Közl.*, 2:93–101, 1954.

[45] Vahid Shahrezaei and Peter S Swain. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105(45):17256–17261, 2008.

[46] Ehud Shapiro, Tamir Biezuner, and Sten Linnarsson. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics*, 14(9):618, 2013.

[47] Josef Stoer and Roland Bulirsch. *Introduction to numerical analysis*, volume 12. Springer Science & Business Media, 2013.

[48] Katjana Tantale, Florian Mueller, Alja Kozulic-Pirher, Annick Lesne, Jean-Marc Victor, Marie-Cécile Robert, Serena Capozi, Racha Chouaib, Volker Bäcker, Julio Mateos-Langerak, et al. A single-molecule view of transcription reveals convoys of rna polymerases and multi-scale bursting. *Nature communications*, 7:12248, 2016.

[49] Mukund Thattai and Alexander Van Oudenaarden. Stochastic gene expression in fluctuating environments. *Genetics*, 167(1):523–530, 2004.

[50] René Thomas and Morten Kaufman. Multistationarity, the basis of cell differentiation and memory. i. structural conditions of multistationarity and other nontrivial behavior. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):170–179, 2001.

[51] NG Van Kampen. *Stochastic processes in physics and chemistry, third edition* . North Holland, 2007.

Model	Monte-Carlo [min]	FD Liouville-master [min]	Push-forward [min]	Gillespie [min]
One gene slow	-	27.5	0.49	6425
One gene fast	-	5.7	0.84	979
Two genes slow-slow	45	-	1.8	15739
Two genes fast-fast	74	-	7	1609
Two genes slow-fast	243	-	6	37935
Two genes fast-slow	249	-	0.76	41232

Table 1: Execution times for different methods. All the methods were implemented in Matlab R2018b running on a single core (multi-threading inactivated) of a Xeon E5 2.4 GHz processor.

Model	State	c_0	c_1	c_2
$G_1 \rightarrow G_2$	(0,0)	f_1	f_2	0
	(1,0)	h_1	f_2	0
	(0,1)	$f_1 + h_2$	0	0
	(1,1)	$h_1 + h_2$	0	0
$G_1 \dashv G_2$	(0,0)	$f_1 + f_2$	0	0
	(1,0)	$f_2 + h_1$	0	0
	(0,1)	f_1	h_2	0
	(1,1)	h_1	h_2	0
$G_1 \rightarrow G_2 \rightarrow G_1$	(0,0)	0	f_2	f_1
	(1,0)	h_1	f_2	0
	(0,1)	h_2	0	f_1
	(1,1)	$h_1 + h_2$	0	0
$G_1 \rightarrow G_2 \dashv G_1$	(0,0)	f_1	f_2	0
	(1,0)	0	f_2	h_1
	(0,1)	$f_1 + h_2$	0	0
	(1,1)	h_2	0	h_1
$G_1 \dashv G_2 \dashv G_1$	(0,0)	$f_1 + f_2$	0	0
	(1,0)	f_2	0	h_1
	(0,1)	f_1	h_2	0
	(1,1)	0	h_2	h_1

Table 2: Constants c_i for computing the next step waiting time in the two gene model.

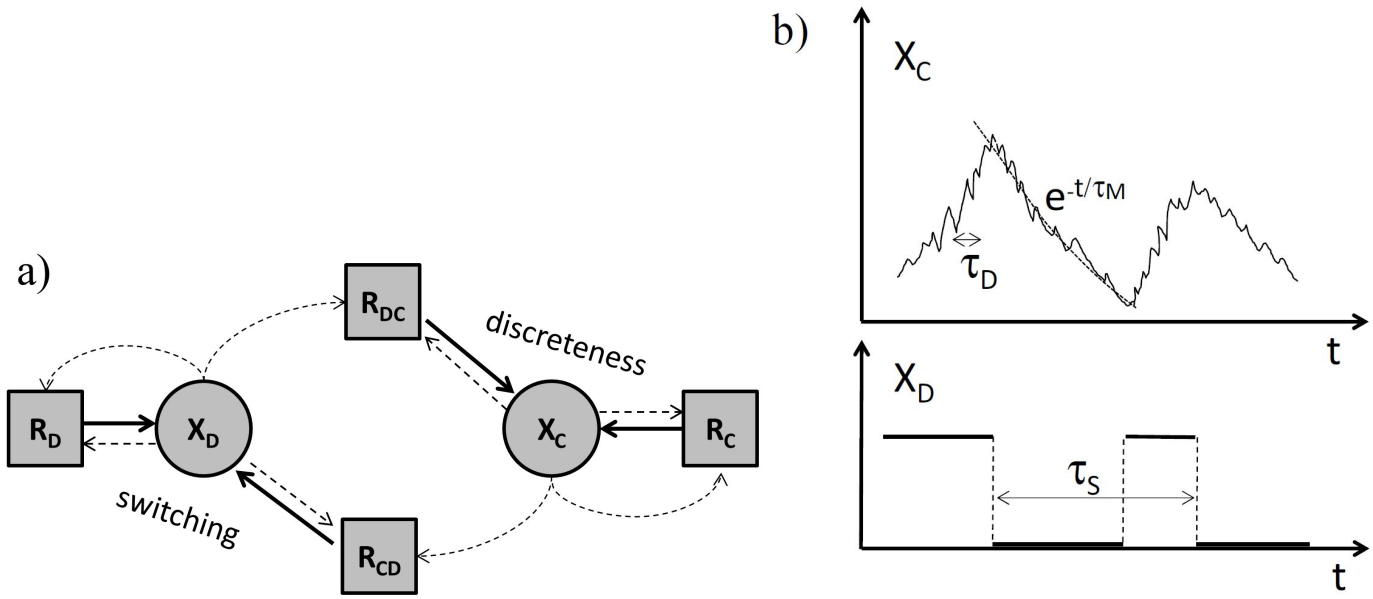


Figure 1: Types of variables and important timescales of stochastic biochemical reactions networks (a) The partition of variables and of the reactions. Dotted line arrows mean that reaction rates depend on the source variables. Continuous line arrow mean that the source reaction acts on the end variable. R_D are reactions acting on discrete variables and whose rates depend on discrete variables, R_C are reactions acting on continuous variables and whose rates depend on continuous variables. R_{CD}, R_{DC} are reactions acting on discrete and on continuous variables, respectively, and have rates depending on both discrete and continuous variables. (b) Typical trajectories and time scales of continuous and discrete variables.

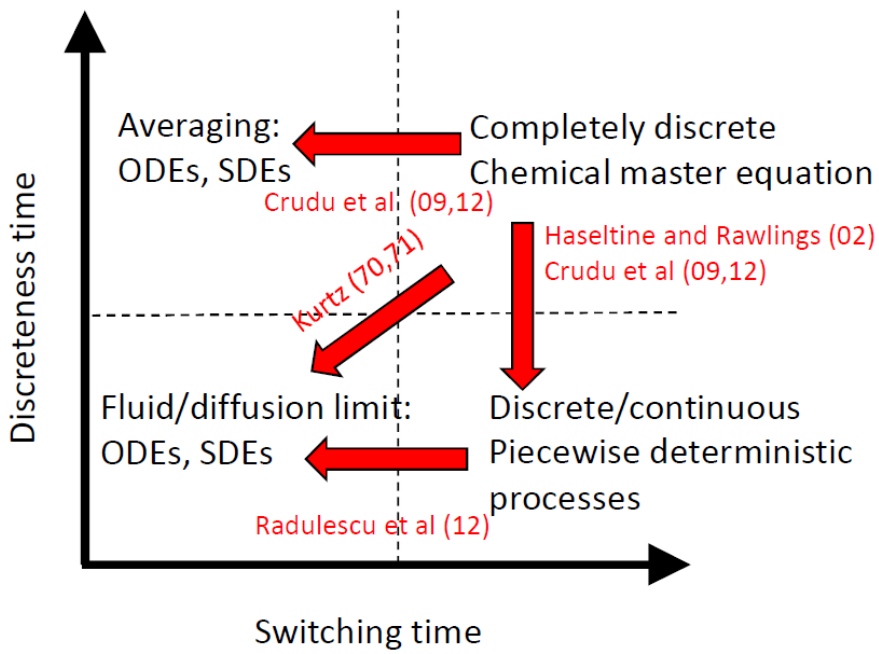


Figure 2: Different approximations of stochastic biochemical reactions networks. The "exact" model is a completely discrete Markov process described by the chemical master equation. Various approximations can be obtained as asymptotic limits when some parameter is very small or very large from the "exact" model or from approximations.

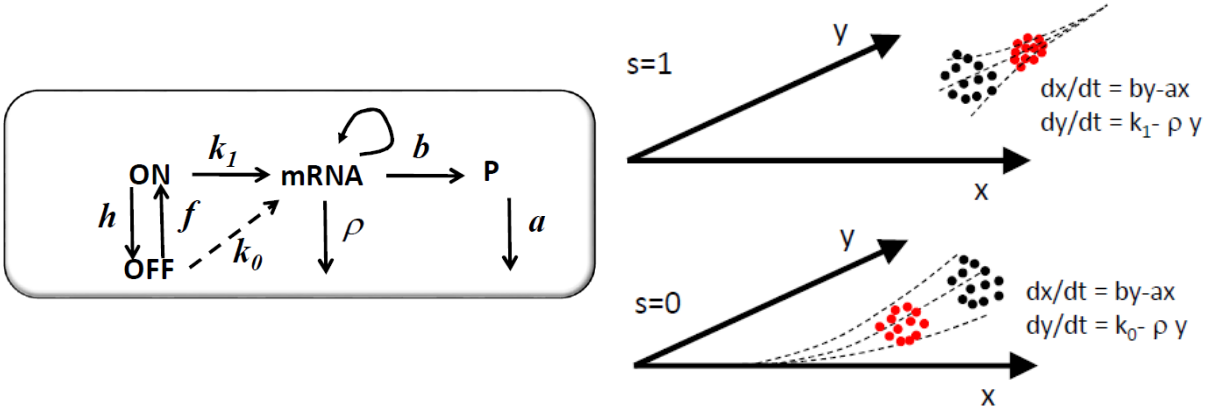


Figure 3: Model of a single gene controlled by a two state (ON/OFF) promoter. In the graphical representation of the model, the arrow from mRNA to mRNA indicates that mRNA is recovered after the transcription reaction. Between two successive jumps of the discrete variables, a probability distribution of the continuous variables x (protein) and y (mRNA) represented as black dots is continuously advected by one of the two possible deterministic flows to a distribution represented by red dots. The $s = 1$ flow pushes the system towards the high expression attractor $x = (bk_1)/(a\rho)$, $y = k_1\rho$ and the $s = 0$ flow pushes the system towards the low expression attractor $x = (bk_0)/(a\rho)$, $y = k_0\rho$, where $k_0 << k_1$. The systems chooses alternately and stochastically between these two flows.

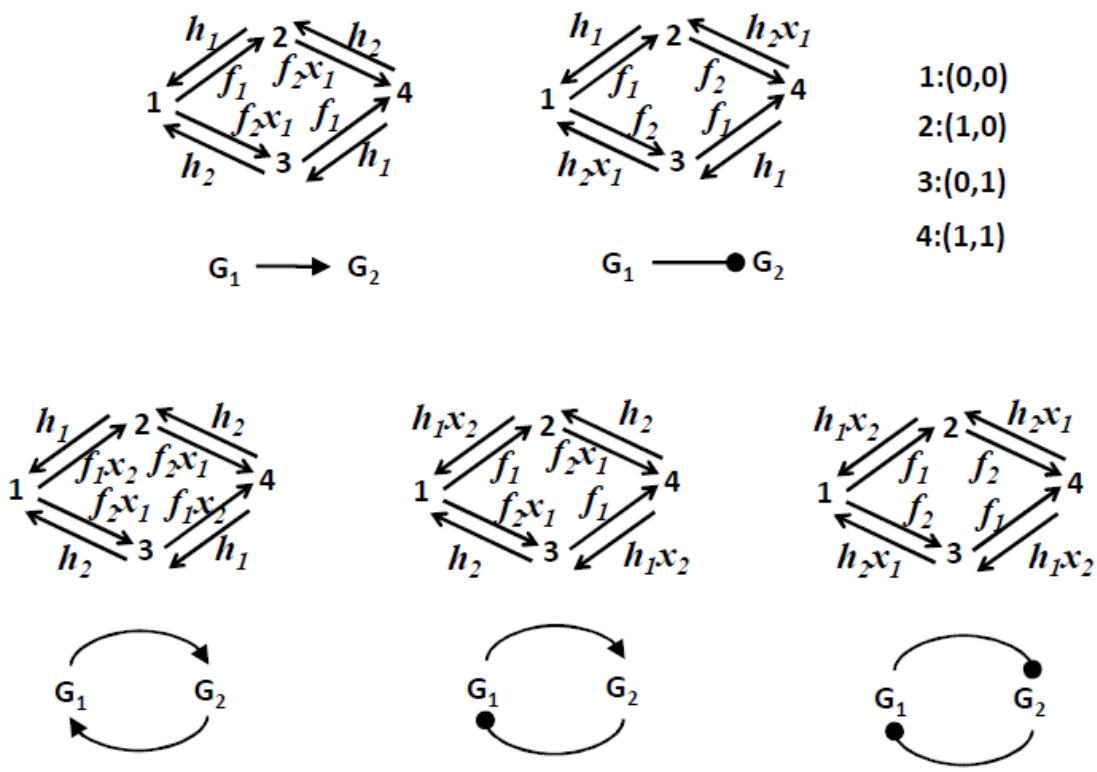
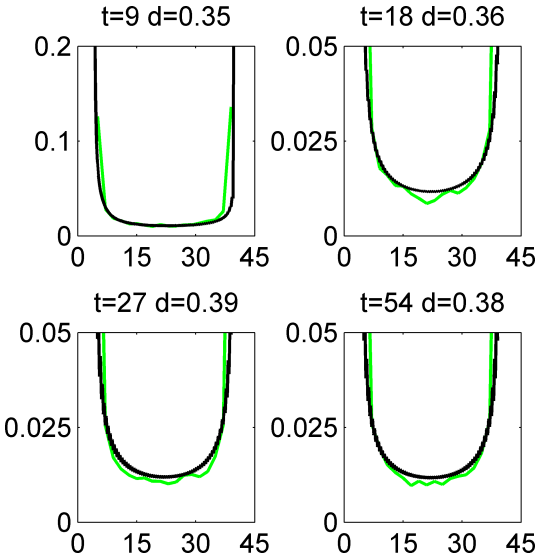
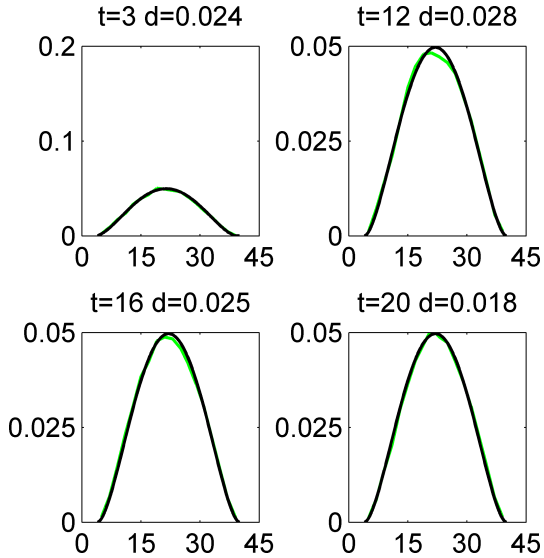


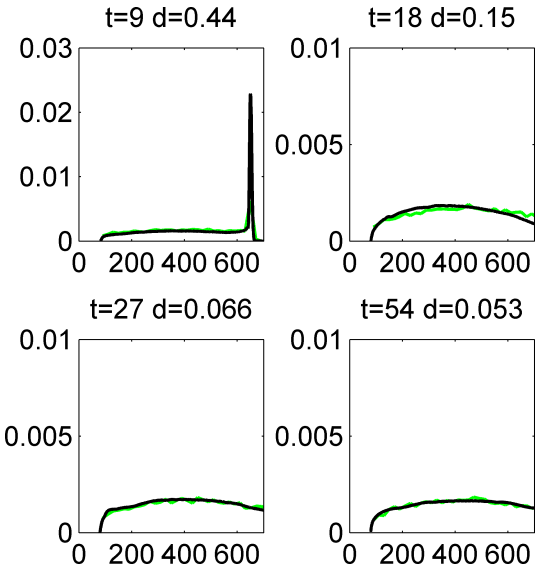
Figure 4: Two gene circuits models. Only the discrete (Markov chain) part of the dynamics is represented. These models have 4 discrete states as each one of the two promoters can be ON or OFF. The transition rates between states are either constant or functions of the levels of proteins x_1 and x_2. We consider that transcription, translation and degradation parameters k_0, k_1, ρ, a, b are the same for the two genes.



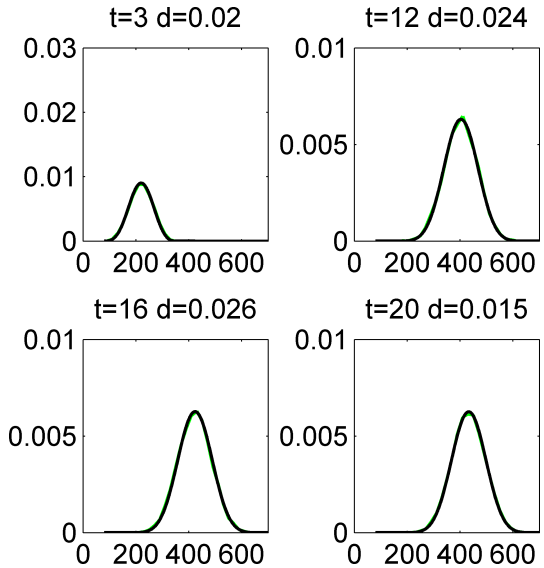
(a) mRNA slow



(b) mRNA fast

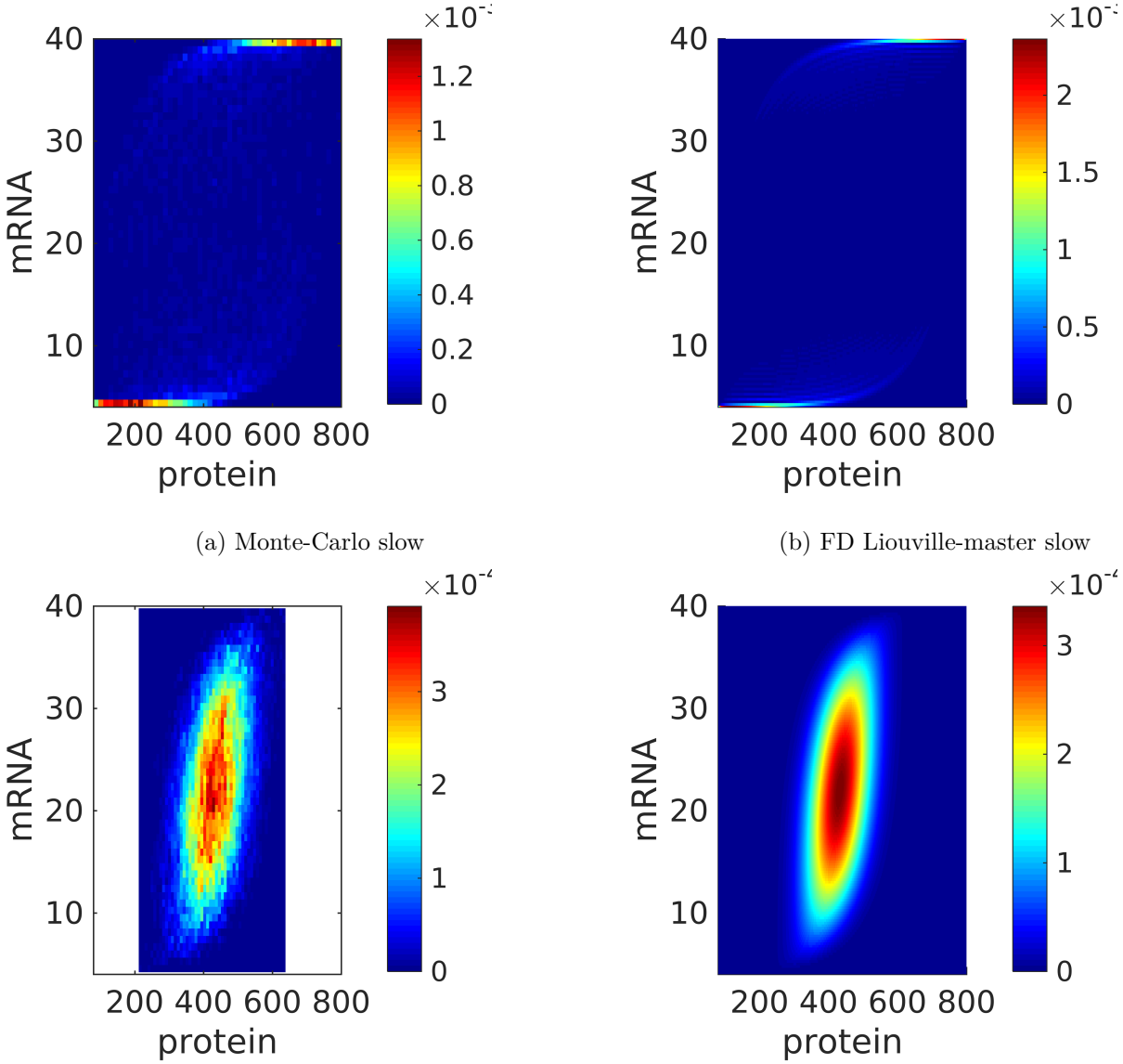


(c) Protein slow



(d) Protein fast

Figure 5: Histograms of protein and mRNA copy numbers for a single gene, produced by the Monte-Carlo method (green lines) and by the finite difference Liouville-master equation method (black lines). The initial data has half-normal bivariate distribution whose mode is $x_1 = x_2 = y_1 = y_2 = 0$. The gene parameters are $\rho = 1$, $k_0 = 4$, $k_1 = 40$, $a = 1/5$, $b = 4$, $p_a = 0.5$, and $\epsilon = 0.5$, $\epsilon = 5.5$, for a slow, and a fast switching gene, respectively. The comparison is quantified by the distance d defined by (4.1).



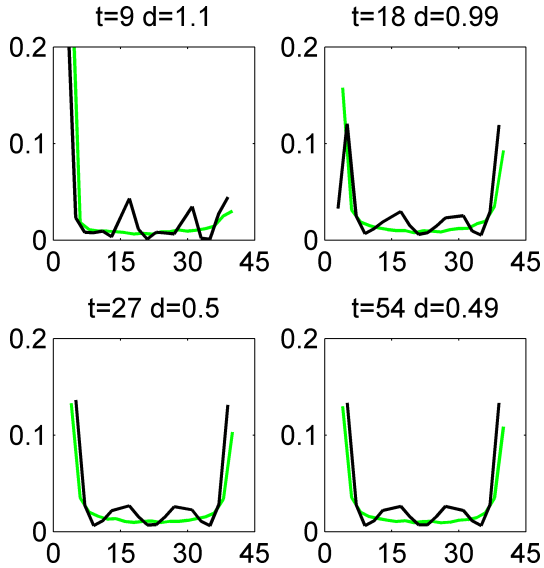
(a) Monte-Carlo slow

(b) FD Liouville-master slow

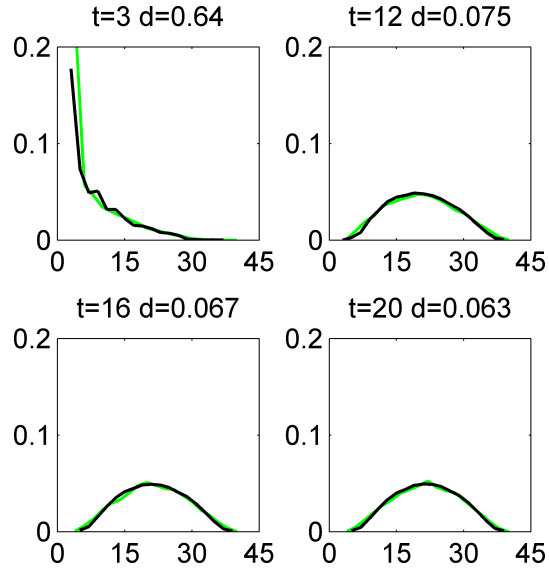
(c) Monte-Carlo fast

(d) FD Liouville-master fast

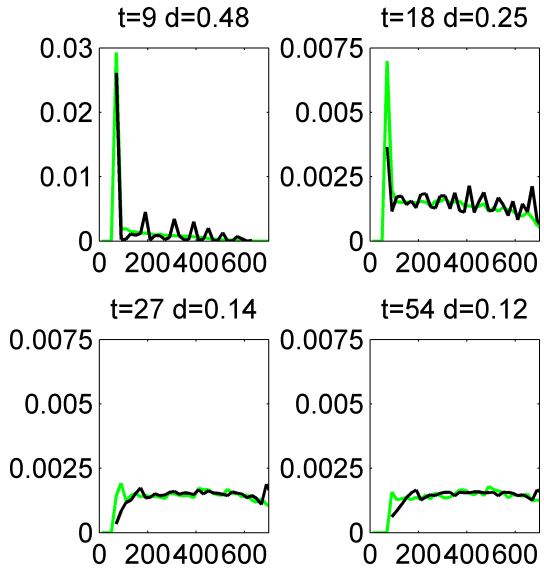
Figure 6: Steady state distributions of protein and mRNA copy numbers for a single gene, produced by the Monte-Carlo method and by the finite difference (FD) Liouville-master equation method. The parameters are those used in Figure 5. The probability to color map relation is linear.



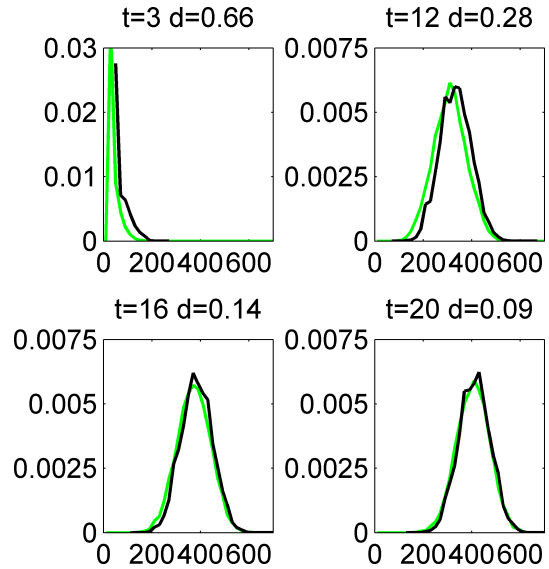
(a) mRNA2 slow-slow



(b) mRNA2 fast-fast



(c) Protein2 slow-slow



(d) Protein2 fast-fast

Figure 7: Histograms of protein and mRNA copy numbers for the second gene in the circuit M1, produced by the Monte-Carlo method (green lines) and by the push-forward method (black lines). The initial data is $x_1 = x_2 = y_1 = y_2 = 0$ and the circuit parameters are $\rho = 1$, $k_0 = 4$, $k_1 = 40$, $a = 1/5$, $b = 4$, $p_a = 0.5$, for both genes, and $\epsilon = 0.5$, $\epsilon = 5.5$, for slow, and fast switching genes, respectively. The comparison is quantified by the distance d defined by (4.1).

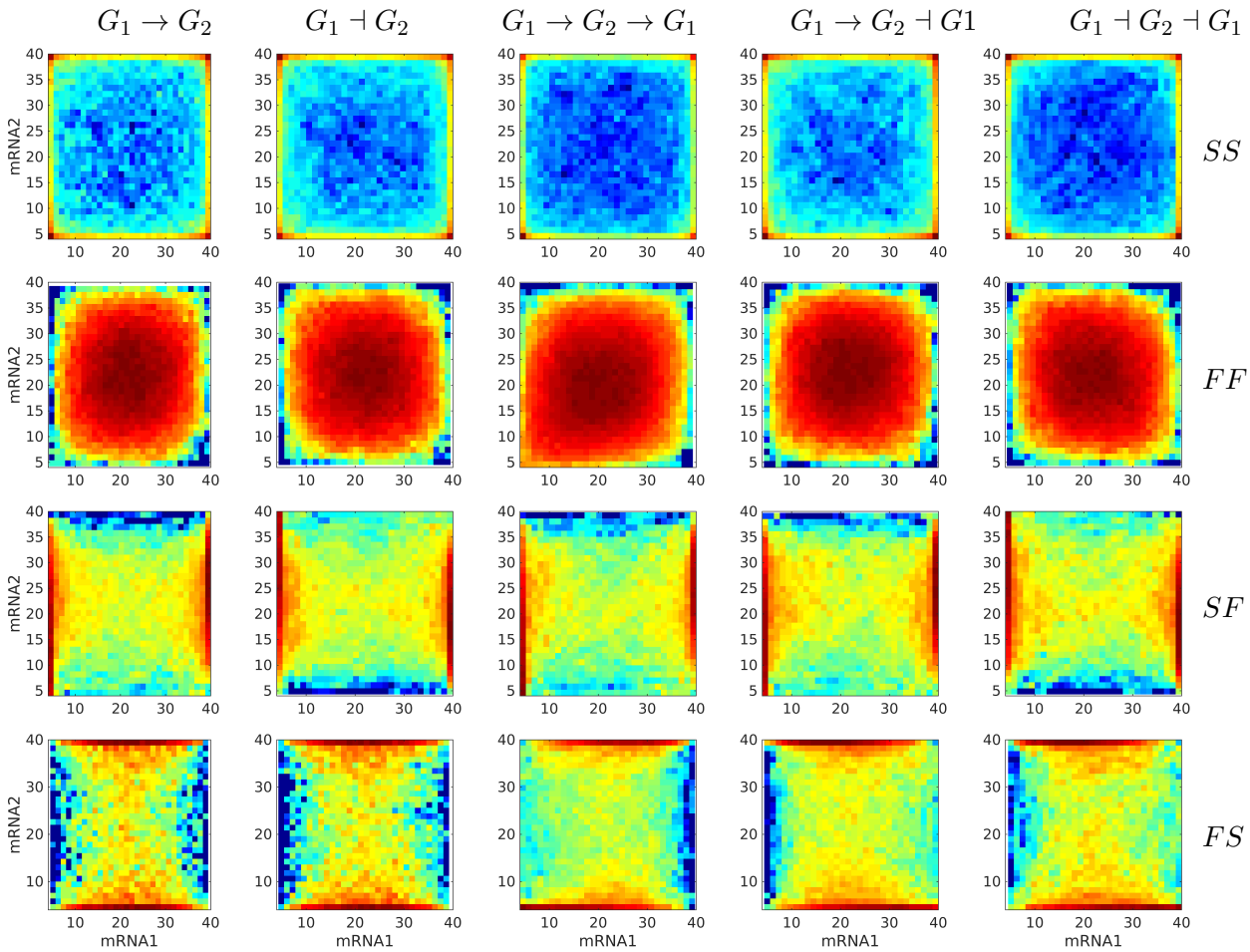


Figure 8: Steady state bivariate histograms of mRNA copy numbers from two interacting genes in circuits of different types and for four switching regimes of the promoters (SS, FF, SF, FS, where SF means that the first gene is slowly switching whereas the second is switching fastly), obtained with the Monte-Carlo method. The individual gene parameters are those used in Figure 5; f and h constants in fx_i or hx_i terms are chosen such that then mean mRNA and protein are the same in regulated and constitutive genes. The probability to color map relation is logarithmic.

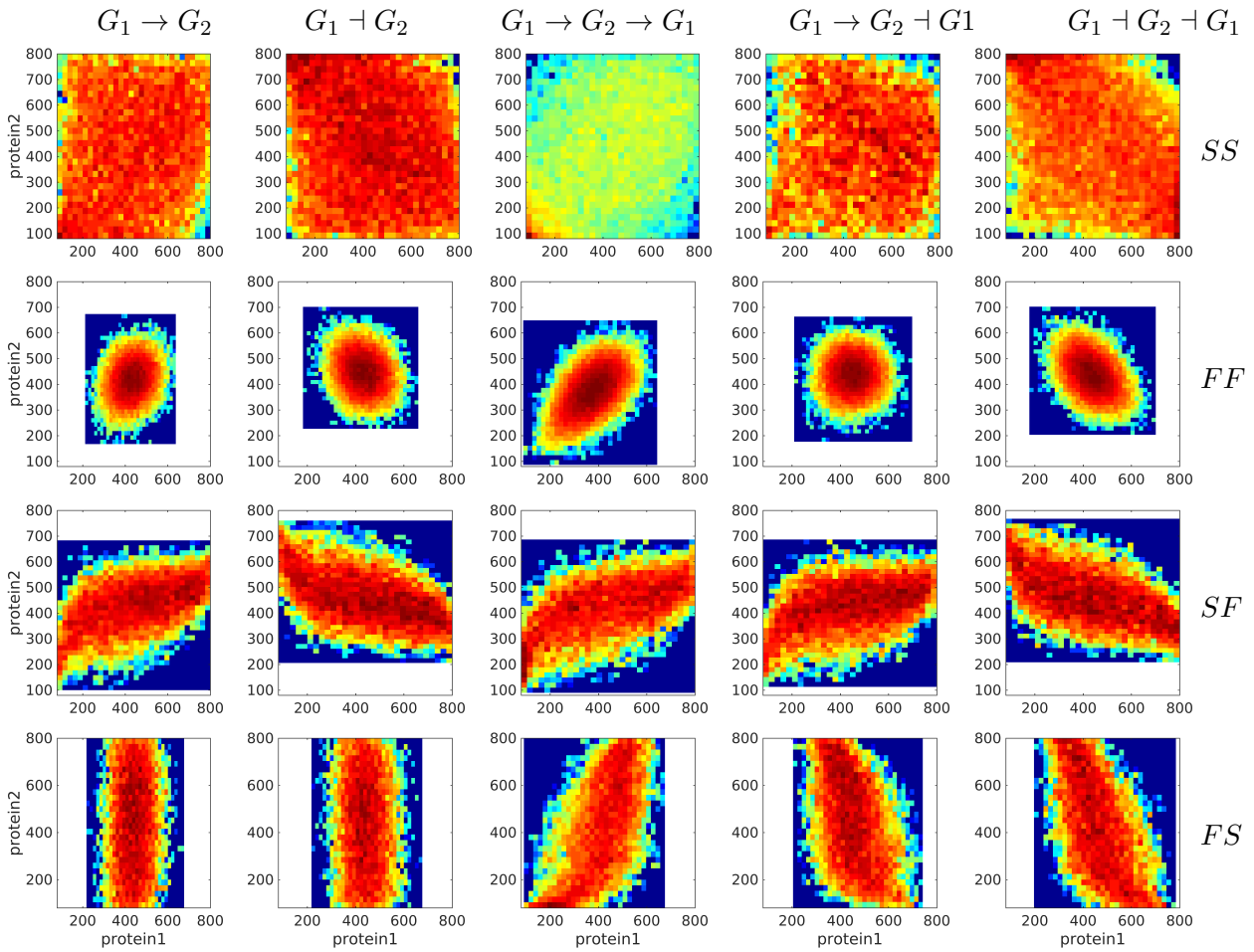


Figure 9: Steady state bivariate histograms of protein copy numbers from two interacting genes in circuits of different types and for four switching regimes of the promoters, obtained with the Monte-Carlo method. The individual gene parameters are those used in Figure 5; f and h constants in fx_i or hx_i terms are chosen such that then mean mRNA and protein are the same in regulated and constitutive genes. The probability to color map relation is logarithmic.

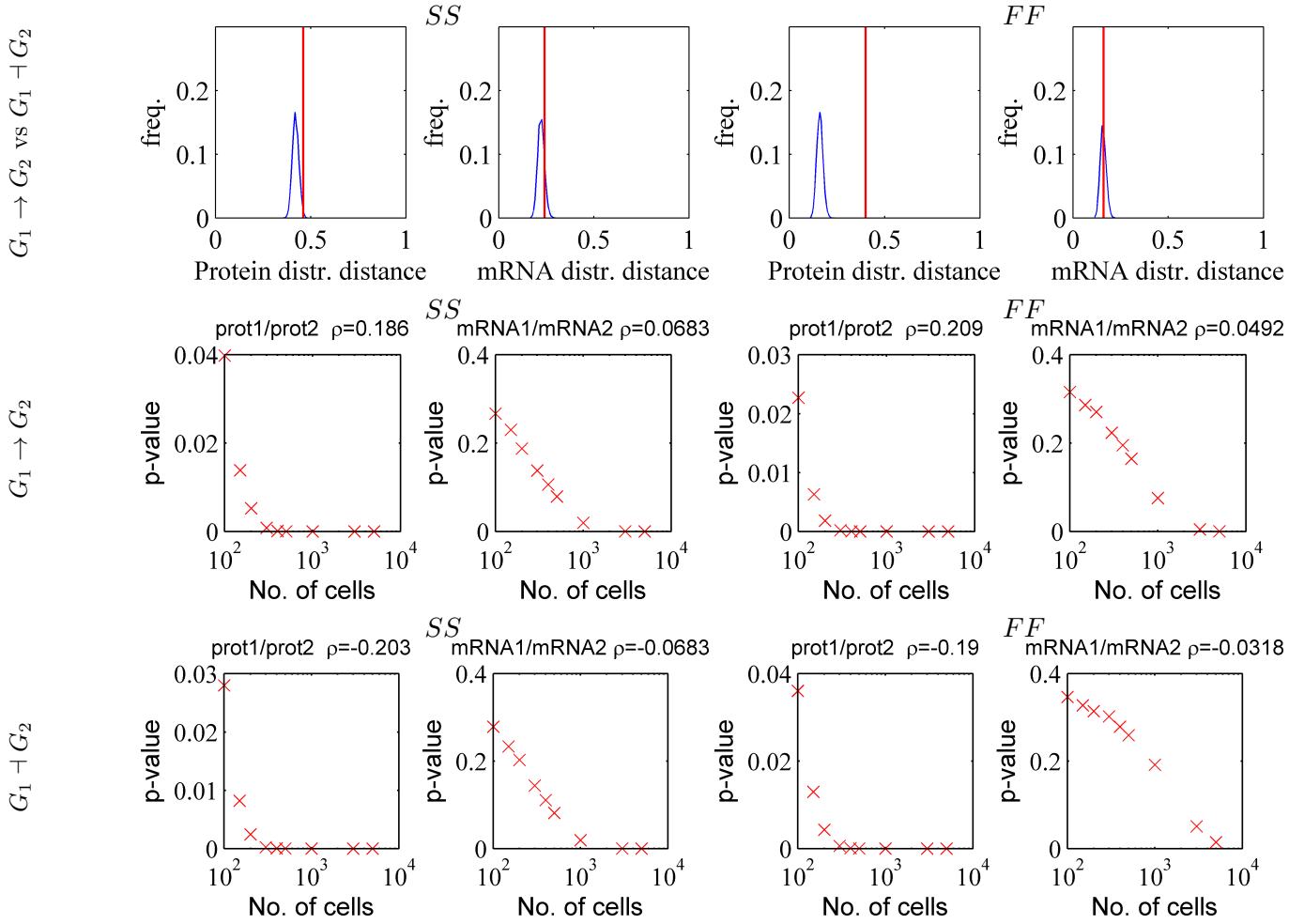


Figure 10: Testing the capacity of mRNA and protein fluctuations to reveal gene/gene interactions. Upper row: the distance between distributions generated by activating and repressive circuit (red line) are compared to the distances between random samples ($N_c = 1000$) generated by the same repressive circuit (blue smoothed histogram). Middle and lower row: the Bravais-Pearson correlation coefficient ρ is computed from random samples containing increasing numbers of cells and the corresponding p-value (p is the upper tail probability and correlation is significant when p is small) is represented vs. sample size. The parameters of the simulations are those used in Figure 9.